DEVELOPMENT RESILIENCE: A COMPARISON OF POPULAR MEASUREMENT METHODS

A Thesis

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Master of Science

by

Susana Constenla Villoslada

August 2019

ABSTRACT

Despite many efforts by governments, international and local non-governmental organizations, and bilateral and multilateral donors, poverty and malnutrition remain prevalent concerns. The need to overcome the dynamic and stochastic nature of poverty and malnutrition to achieve long-term development goals has brought attention to the role that shocks and stressors play in determining well-being outcomes. Resilience, the capacity over time of avoiding poverty, however defined, in the face of multiple shocks and stressors, picks up this concern. The concept has increased in popularity in development circles during the last decade, but, to date, no consensus has been reached on how to measure it empirically. This paper carries out a comparative analysis of three of the most currently popular resilience measurement methods, with the ultimate goal of determining how similar they are to one another and to standard well-being measures. I find important differences in how the models measure households' resilience capacity, and how they relate to current and future well-being. Due to differences in how they rank households in terms of resilience capacity and how they identify resilient units, their results are not directly comparable. The same households do not get consistently high or low resilience capacity values through all methods, and in two of the three, resilience capacity measures are very dissociated from standard well-being outcomes, both cross-sectionally and in inter-temporal well-being prediction.

BIOGRAPHICAL SKETCH

Susana Constenla-Villoslada was born in A Coruna, a coastal region in the North-West of Spain. She started her post-secondary school academic journey with an advanced vocational training diploma in agricultural businesses management. She then moved to Madrid, where she obtained her degree in Agricultural Engineering by the Polytechnic University of Madrid. Thanks to the support of a Fulbright Scholarship, she was able to continue her academic training in the US with a Master of Science degree in Applied Economics and Management at Cornell University.

To my close family, that have always let me do me.

ACKNOWLEDGEMENTS

I want to express great gratitude to my chair, professor Christopher B. Barrett, for giving me the opportunity of working on this project. He did not know me from before, but still he gave me the trust to work with me. I was left out without a thesis because of issues with another project, and, even though it was already late to start something new, he gave me the opportunity to prove my self and try. I feel very honored and thankful for having had the chance to work with such an exceptional advisor, professional and human being in all aspects.

I also want to thank my minor chair, professor John F. Hoddinott, to make some time to give me his advise on this project. Professor Hoddinott is also a great example of moral integrity and exceptional professionality. Having his advise through these months has been incredibly useful.

I am also indebted to Joanna Upton, whose previous work on the topic and joint meetings and discussions allowed me to work faster and think better on this project. She has proved herself as a great person to work with, always ready to help and sincerely preoccupied about the well-being of the ones that surround her.

Finally, I want to thank my mother and father, Clafira and Manuel, for all the support that they have given me to continue with my academic training. I would not be here without them. I also want to thank my sister, Cristina, for being always there for me and reminding me of who I am in moments of doubt. I also want to thank my aunt, Angeles, for supporting me in the toughest moments; sometimes a hug worths more than a 1,000 words. Lastly, my friend Alex has to be in these lines, because he has been present trough almost all of my life's journey.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Despite many efforts by governments, international and local non-governmental organizations, and bilateral and multilateral donors, malnutrition and poverty remain widespread concerns partly because their dynamic and stochastic nature makes it difficult to achieve development goals in the medium and long run. Those living below acceptable well-being levels are commonly threatened by recurrent climate, economic, conflict and other shocks. Research on the causes of vulnerability shows that the impact of those and other shocks is not uniform, rather is often aggravated or ameliorated by some characteristics of those households and individuals exposed to them (Naud'e, Santos-Paulino and McGillivray, 2009). Resilience goes further than vulnerability, bringing to the table the role that shocks and risk play in reaching development goals with the ultimate objective of identifying and strengthening the characteristics that lead to shock-proof communities whose well-being can increase and remain stable over time. Formally, resilience in development has been defined by Barrett and Constas (2014) as the capacity over time of avoiding poverty, however defined, in the face of multiple shocks and stressors, and as the capacity that ensures adverse stressors and shocks do not have long-lasting adverse development consequences, by Constas, Frankenberger and Hoddinott (2014).

In order to strengthen resilience capacity, we need to be able to measure it empirically. Since the pioneering work of Alinovi et al. (2008) under the Food and Agriculture Organization of the United Nations (FAO), there have been multiple attempts to generate resilience measurement tools that can be implemented in the field. Those methods differ, both in theoretical conceptualization and statistical approach. To date, no agreement has been reached to generalize one particular measuring procedure. Nevertheless, the pressing need of development agencies to implement resilience capacity building interventions has made them adopt some of those methods as official, making them common practice.

Three resilience measuring tools stand out from the others for their relatively widespread use. The FAO and other United Nations affiliate agencies (e.g., WFP, IFAD, etc.) use the Resilience Indicators for Measurement and Analysis (RIMA) as their resilience measurement tool. Originally developed by Alinovi et al. (2008, 2010), and recently updated by the FAO under the RIMA-II framework (from hereon RIMA) (FAO, 2016), the method constitutes the first and perhaps the most widespread quantitative tool to measure resilience capacity empirically. Under RIMA, four factor analysis indexes, called ”pillars” (Access to Basic Services (ABS), Assets (AS), Social Safety Nets (SSN), and Adaptive Capacity (AC)), operationalize resilience capacity. Under a Structural Equation Model-Multiple Indicators Multiple Causes (SEM-MIMIC) framework, a resilience capacity index (RCI) is generated. The RCI is conceptualized as the latent variable mediating the effects between the pillars and some food security indicators of interest (e.g., Food Consumption Score (FCS)). The variable selection to generate the pillars is informed by previous work in the resilience, vulnerability and food security literatures, and kept as simple as possible to ease the implementation of the method with standard community and household level data, much like that in the Living Standards Measurement Surveys (LSMS), which I use in this thesis. To classify units of analysis as resilient or not, RIMA does not use the estimated RCI, but instead uses a binary classification method based on the dynamic analysis of well-being between subsequent time periods. Under RIMA, units that have managed to not suffer a loss in well-being after being affected by a shock are classified as resilient. Those who had a decline in well-being between survey rounds are classified ”not resilient”. However, it is important to clarify that RIMA does not attempt to use this classification method as a targeting procedure. They refer to it more as an indirect identification method for resilience.

TANGO international developed another resilience measurement method (Frankenberger and Smith, 2016, Smith and Frankenberger, 2018; TANGO from hereon) that was chosen by the United

States Agency for International Development (USAID) for resilience analysis (Henly-Shapard and Sagara, 2018). Similar to RIMA, under this method resilience capacity is also conceptualized as a latent variable. However, TANGO international follows the original structure proposed by the Food Security Information Network (FSIN) (Constas, Frankenberger and Hoddinott, 2014; Constas et al. 2014), estimating resilience capacity as a factor analysis index that combines absorptive, adaptive and transformative capacities. The last three are also factor analysis indexes, generated from a wide range of complex indicators, several of them also indexes constructed using factor analysis. The binary classification of the population as resilient or not is done in the same way as in the RIMA case, by identifying units that have not suffered a loss in well-being after being affected by a shock.

Cissé and Barrett (2018, hereafter CB) authored the last of the three methods studied here. Generated and used predominantly in academic research (e.g., Upton, Cisse and Barrett 2016, Cisse and Ikegami 2018, Alloush 2019, Knippenberg et al. 2019, Phadera et al. 2019), this method is the most distinguishable from the other two. Here resilience capacity is not conceptualized as a latent variable or factor index, but as the conditional probability of being above a certain normative well-being threshold at a specific moment in time. Under CB, a sequence of two ordinary least squares (OLS) regressions are estimated to generate the conditional mean and variance of a conditional well-being probability density function (p.d.f.), from which one can readily compute the probability of being above the pre-specified threshold. That conditional probability represents the resilience score (RS). In the case of CB, the identification of resilience is directly tied to the resilience score. In CB, units are resilient if and only if they have a resilience score higher than a pre-specified probability threshold (e.g., probability of not being poor higher than 40%).

These three methods differ in their core features, such as what characteristics build resilience capacity, the statistical approach to identify and measure it, and how to define a unit as resilient or not. However, no comparative analysis of the results derived from the use of each method has been done yet. Consequently, it is not clear to what extent these three measures are related to one another or to standard well-being measures. This paper fills this gap, with the joint estimation of RIMA, TANGO and CB methods with a common dataset, so as to be able to perform a comparative analysis between their resilience measures. This paper by no means pretends to offer an exhaustive analysis of all the resilience measurement tools available. The selection is based on the widespread use of those three measures in important development circles, and on the possibility of estimating them with a common panel dataset, which is essential for comparative purposes. Following as closely as possible the guidelines proposed by the authors in their original papers, we replicate the three resilience capacity measures, and perform a set of comparative tests to see to what extent those measures correspond with each other and to standard well-being measures. The comparative analysis that we carry out is designed to determine how similar do the methods classify the sample in terms of resilience capacity; the internal consistency of the methods regarding resilience capacity measures and binary classifications of resilience; their out-of-sample well-being predictive accuracy; and the buffering capacity of the resilience capacity measure in the face of shocks, measured as exposure to drought in our study. To clarify, the comparative analysis is designed to answer the following questions:

1. How similarly does each method classify the sample in terms of resilience capacity? Are the same households getting similar RCI/RS values in each of the methods?

2. Are both resilience measures, the resilience capacity and the binary resilience indicator, consistent between each other within each method? Do resilient units have higher estimated resilience capacity?

3. How does the binary classification of resilience relate to future normative well-being status? Are resilient units more likely to satisfy normative standards of well-being in the future?

4. How accurately does resilience capacity predict well-being out-of-sample, in either cross-section or in future periods?

5. Do the resilience capacity measures differ on how they mediate the negative effects of shocks on well-being?

The comparative analysis of the non-parametric densities and Spearman rank correlation coefficients indicate that RIMA and TANGO resilience capacity indexes are reasonably highly correlated between each other, but much less with the CB resilience score. Conversely, the CB resilience score correlates better with standard well-being measures of poverty and food security than do the RIMA and TANGO measures. RIMA and TANGO yield similar measures of resilience capacity, but CB's resilience score is a closer representation of actual well-being.

The internal consistency checks show that resilient and non-resilient groups under the TANGO/RIMA binary classification do not have significantly different average resilience capacity values. This is perplexing and arises a serious issue about the usefulness of either or both of their indicators since they are inconsistent with each other for both methods. In the CB case, the resilient and non-resilient groups have significantly different resilience scores, which is in line with the fact that the binary classification is directly tied to the resilience capacity measure.

In terms of out-of-sample prediction of cross-sectional and future well-being, CB outperforms both RIMA and TANGO. The result holds when well-being is predicted by the resilience capacity measure and also by the main statistical model that relates resilience capacity and well-being outcomes.

Finally, in our sample, in all cases but one the resilience capacity measures do not have a significant estimated effect on buffering the negative effects of drought. Nevertheless, the limitations of the data should be taken in account when interpreting the results of that analysis. Because we cannot know the true value of the drought-buffering role played by resilience capacity, we do not have a benchmark against which to judge any of the three measures. We can only observe that one measure, TANGO, generates a regression coefficient estimate of the sign one would expect conceptually, although the true sign in these data is unknown.

The rest of the paper is organized as follows. Section 2 provides an overview of the theoretical framework and estimation procedure of the three methods, including the main empirical applications for which they have been used. In section 3 we present the empirical strategy of the replication and comparison exercises. Section 4 shows the data used and variables generated, and section 5 the results, both descriptive and comparative. Finally, in section 6 we include some reflections and recommendations.

CHAPTER 2

RESILIENCE MEASUREMENT METHODS

2.1 Resilience Index for Measurement and Analysis (RIMA)-II

Starting with the work of Alinovi et al (2008, 2010), the FAO pioneered the creation of both a theoretical framework and a quantitative empirical approach for the analysis of development resilience. That first quantitative measurement tool was the Resilience Index for Measurement and Analysis (RIMA), which became the most widely used tool for guiding and evaluating interventions designed to build resilience capacity. Incorporating the knowledge accumulated in its years of application and the recommendations of FSIN (d' Errico, Garbero and Constas, 2016; Constas, Frankenberger & Hoddinott, 2014; Constas et al., 2014), RIMA was revised and re-released in 2016 under the RIMA-II framework (FAO, 2016). The method has been used for impact evaluations of resilience capacity building interventions (Garbero and Chichaibelu, 2016) and the empirical identification of the relationship between resilience capacity and changes in food security between pre and post-shock periods (dErrico et al., 2018).

Under RIMA, resilience is the capacity that makes sure that shocks and stressors do not have long-lasting development consequences. That capacity is built by a set of characteristics at different aggregation levels - community and household - the selection of which is based on previous research on resilience, vulnerability and food security. The variables are grouped into four factor analysis indexes, called "pillars" -Access to Basic Services (ABS), Assets (AS), Social Safety Nets (SSN) and Adaptive Capacity (AC)-. Each of the four pillars is designed to capture a different dimension of resilience (refer to Appendix A for a detailed explanation). A fifth pillar that captures shock exposure, Sensitivity (S), is mentioned in the original document but was not included in the direct estimation of the resilience capacity measure. Under a SEM-MIMIC framework, a resilience capacity index (RCI) is estimated as a latent variable mediating the effect between the pillars and some well-being outcomes of interest. Mathematically, the model is specified as follows,

$$[RCI]_{it} = [\beta_1, \beta_2, ..., \beta_n] * [ABS, AST, SSN, AC]_{it} + [\epsilon_1] \quad (1)$$

and

$$[W_1, W_2, .., W_n] = [\alpha_1, \alpha_2, ..., \alpha_n] * [RCI] + [\epsilon_2, \epsilon_3, ..., \epsilon_n] \quad (2)$$

where RCI is the resilience capacity index of household i at time t, ABS, AST, SSN, AC are the four factor index pillars and W_n are different well-being outcomes. Both equations are solved simultaneously under the SEM-MIMIC, which allows for the prediction of the RCI. The final RCI is a unitless, numerical index, specific for each household in the sample, and re-scaled by a mini-max procedure to range between 0 and 100.

The purpose of the RCI is to measure resilience capacity, but the classification of households as resilient or not is not done using this unitless index. Under the RIMA method, resilient are households that, after being affected by a shock, do not register a loss in well-being between two consecutive time-periods. Consequently, the resilience classification of households is completely dissociated from the resilience capacity measure itself. Mathematically, R is the binary resilience classifier such that,

$$R = \begin{matrix} 1 \; if \; Y_t \geq Y_{t-1} \\ 0 \; if \; Y_t < Y_{t-1} \end{matrix} \quad (3)$$

Nevertheless, the FAO does not use binary definition of resilience for targeting purposes, but rather to explore the relationship between resilience capacity and resilience, defined as nondeclining measure of well-being. Through a probit model, the probability of being non-resilient, or suffering a loss in well-being, is explained by the variables that build the resilience pillars or by the RCI itself (FAO, 2016; d'Errico et al., 2018). Since the probit specification is not directly comparable with the other two methods, we set it aside and focus on the estimation of the RCI and of the binary resilience classification.

It is important to note that measures of shock exposure are not directly included in any of the previous equations. In their original paper, RIMA uses the measures of shock exposure just to determine the households that were affected by a shock and suffered a loss in well-being or not and, in a separate model, to estimate the effects of the shocks on the RCI. Consequently, the impact of shocks on well-being is not analyzed and the relationship between shocks, resilience capacity and wellbeing is not formalized in an integrated mode. However, d'Errico et al. (2018) does estimate a probit model with the probability of suffering a loss in well-being as the dependent variable that includes both the RCI and measures of shock exposure. The buffering capacity of the RCI on the negative effects of shocks is therefore directly tested in this revised application of the RIMA method.

2.2 The TANGO method

Building on the RIMA method and on the work of the FSIN (Constas, Frankenberger and Hoddinott, 2014; Constas et al., 2014), Frankenberger and Smith (2016) and Smith and Frankenberger (2018), from the development consulting firm TANGO, developed an empirical resilience measurement tool (TANGO hereon) which the main purpose of which is the analysis, in a regression framework, of the relationship between well-being and resilience capacity, in particular how resilience capacities mediate the effects between well-being and shocks. Since 2018, TANGOs method has become the official resilience measurement tool of USAID funded projects (Henly-Shapard and Sagara, 2018).

The TANGO method has multiple similarities with RIMA. For TANGO, resilience is also the capacity that ensures that shocks and stressors do not have long-lasting development consequences.

Resilience capacity is also conceptualized here as a latent variable, estimated by factor analysis from the combination of three other indexes constructed by factor analysis. Those indexes are the absorptive, adaptive and transformative capacities proposed in the FSIN resilience guidelines. They are built from an extensive set of variables, the complexity of which surpasses that of RIMA's pillars (for a detailed explanation of the dimensions of resilience that each of the three capacities tries to capture, refer to Appendix A). The combination of the absorptive, adaptive and transformative capacities through factor analysis yields, as in RIMA, a unitless resilience capacity index re-scaled by a mini-max procedure to a 0-100 range.

The RIMA and TANGO methods also share the way on which they classify units as resilient or not. For TANGO, resilient units do not suffer a loss in well-being between two successive time periods after being affected by a shock, (equation 3, pg.7). Nevertheless, it is important to clarify that, as in the RIMA case, the authors do not at any point advocate for the use of that classification for targeting purposes.

In TANGO's original paper, the resilience capacity measure is mainly used in a regression framework, to test its relationship with a well-being outcome in the face of shocks. The authors propose two main specifications: a cross-sectional OLS regression and a panel growth model. Mathematically,

$$Y_{it} = \alpha + \beta_1 SE_{it} + \beta_2 RCI_{it} + \beta_3 X_{it} + \epsilon_{it} \tag{4}$$

$$Y_{it} - Y_{it-1} = \alpha + \gamma_1 SE_{it-1} + \gamma_2 RCI_{it-1} + \gamma_3 X_{it} + u_{it} \tag{5}$$

where Y is a measure of well-being of household i at time t, SE refers to shock exposure between $t-1$ and t, RCI is the resilience capacity factor index and X_i a vector of household characteristics. To determine the mediating effect of the RCI between well-being and shocks, equations 4 and 5 are slightly modified with the incorporation of an interaction term between the RCI and the measure of shock exposure,

$$Y_{it} = \alpha + \beta_1 SE_i + \beta_2 RC_{it} + \gamma SE_i * RCI_{it} + \beta_3 X_{it} + \mu_v + \epsilon_{it} \tag{6}$$

$$Y_{it} - Y_{it-1} = \alpha + \beta_1 SE_i + \beta_2 RCI_{it-1} + \gamma SE_i * RCI_{it-1} + \beta_3 X_{it} + \mu_v + \epsilon_{it} \tag{7}$$

For the purpose of our study, we are mostly interested in the estimation of the resilience measure and the binary classification of resilience, since those are directly comparable among methods, and will not estimate the other models proposed by TANGO. However, we use also a modified version equation 6 to test the buffering capacity of the negative effects of shocks on well-being of the three resilience measures.

When estimating factor indexes, TANGO uses a modified version of the standard principal factor analysis, which we will also use in an attempt to follow as close as possible their original guidelines. In their version, just the first wave is used to generate the factor loadings that are then applied to normalized versions of the variables from other waves. Best practice factor analysis uses the data from all of the survey waves to which the analysis will be applied so as to account for stochasticity in the variable measures (due both to true stochasticity and measurement error). TANGO also employs an unconventional, manual screening process of the variables to include in the final factor. Variables that have the ”wrong sign” according to the author’s theoretical interpretation, and/or which result of the Kaiser-Meyer-Olkin (KMO) test is lower than 0.50 are discarded from the estimation of the final factor. These are unorthodox practices in the factor analysis literature, but we nonetheless follow them in replicating the TANGO analysis so as to maintain fidelity to the original authors’ approach. A detailed explanation step by step of this procedure can be found in Appendix B. Lastly, when estimating factor analysis indexes, TANGO includes repeated variables in different indexes, which is also an unorthodox practice but that we also follow in an attempt to replicate as close as possible the author’s original guidelines. The variables that constitute each index will be detailed in later sections of this thesis.

2.3 Cissé and Barrett and the stochastic dynamics of well-being

Cissé and Barrett offer a quantitative measure of resilience that is the empirical implementation of Barrett and Constas (2014)'s theoretical conceptualization of development resilience. Cissé and Barrett (2018) use data from northern Kenya to illustrate the method, and offer an illustrative comparison of the resilience scores of different sub-populations. Others studies have combined the method with a rigorous causal identification strategy for impact evaluation. For example, Cissé and Ikegami (2016) assess the impacts of Index Based Livestock Insurance on pastoralists' resilience in northern Kenya, and Phadera et al. (2019) assess the impact of livestock transfers on households in Zambia. Alloush (2019) estimates resilience using panel data from South Africa and relates it to psychological well-being, while Knippenberg et al. (2019) apply the method to high frequency data from rural Malawi. Upton et al. (2016) explain how the CB methodology, when applied to food security indicators such as FCS, food expenditures, or dietary diversity, also provides a food security measure that more closely matches the internationally agreed (1996 World Food Summit) definition of food security than do existing food security measures.

In CB, resilience is defined as the capacity that allows the unit of analysis to avoid poverty, however defined, over time, even in the face of various shocks and stressors. Their methodology is not based on the estimation of a latent resilience capacity index similar to the ones of TANGO or RIMA. Rather, CB identifies resilience through what the authors refer to as the resilience score, which corresponds to the conditional probability of being above a pre-specified normative well-being threshold. That probability is unique for each unit of analysis, derived from the empirical estimation of a unit-specific probability density function (p.d.f) of well-being. The functional form of the p.d.f is imposed by the researcher, but CB suggests the use of normal, lognormal or gamma distributions. The conditional mean and conditional variance to build the conditional probability distribution are generated through the estimation of the following equations,

$$\hat{\mu}_{it} = \hat{E}[W_{it}|W_{it-1}, X_{it}] = g_M(W_{it-1}, X_{it}, \hat{\beta}_M) + \hat{\delta}_{1M} X_{it} + \hat{\delta}_{2M} C_{it} + \hat{\gamma}_M S_{it} + u_{Mit} \tag{8}$$

$$u^2_{Mit} = \hat{\sigma}^2_{it} = g_V(W_{it-1}, X_{it}, \hat{\beta}_V) + \hat{\delta}_{1V} X_{it} + \hat{\delta}_{2V} C_{it} + \hat{\gamma}_V S_{it} + u_{Vit} \tag{9}$$

Equation 8 yields the mean, $\hat{\mu}_{it}$, of the well-being p.d.f for household i at time t. The function $g(.)$ refers to a polynomial function of lagged well-being $W_{i,t-1}$, $X_{i,t}$ a set of household characteristics, $S_{i,t}$ a measure of shock exposure and $C_{i,t}$ a set of community characteristics. Equation 9 estimates the conditional variance, $\hat{\sigma}^2_{it}$, where the dependent variable are squared the residuals, u_{Mit}, from equation 8. The first equation is estimated through regular ordinary least squares (OLS), while the second equation is fitted by generalized linear models, in particular Poisson (log-linear). The CB method is less strict in determining what builds resilience capacity. The selection of the household and community explanatory variables and controls is left to the researcher's criteria.

With those two conditional probability moments (mean and variance) of the conditional well-being p.d.f., a resilience score can be derived for each household. The resilience score, the conditional probability of being above a normative well-being threshold, is formalized mathematically as follows. For each time period s up to T periods into the future we construct the resilience score $(\rho_{i,s})^T_{s=0}$ where:

$$\rho_{i,s} \equiv Pr_{i,t+s} \geq (\underline{W}|W_{it}, X_{it}, S_{it}, C_{it}) = F(\underline{W}, W_{it}, X_{it}, S_{it}, C_{it}) \tag{10}$$

where $F(.)$ is an assumed two-parameter inverse cumulative density function. We use the gamma distribution following Cissé and Barrett (2018). Contrary to the resilience capacity measure in RIMA and TANGO, which is unitless, in CB it represents a probability measure, which therefore falls between a 0-1 range. However, by multiplying it by 100, we can obtain a measure that ranges between 0-100 and can be directly compared to the resilience measures of the other two methods.

In CB, the identification of the units as resilient or not is directly related with the generation of the resilience score. Resilient are the households that have a resilience score, $\rho_{i,s}$, higher than a pre-specified probability threshold. As with the well-being threshold itself, this probability threshold may vary based on the goals of the assessment. We may, for example, consider those resilient who have a relatively low probability of falling above a relatively high poverty threshold; whereas, if we are concerned about humanitarian emergencies we may prefer to consider those as resilient who have a very high probability of avoiding falling below a very low threshold. For the purposes of direct comparison with RIMA and TANGO in the analyses that follow, we designate as resilient the same percentage of households that corresponds with the percentage identified by the RIMA/TANGO binary resilience indicator defined in equation 3.

CHAPTER 3

EMPIRICAL STRATEGY- COMPARATIVE METHODS

Our empirical approach is based on the comparison of the resilience capacity indexes (RCI) in RIMA and TANGO and the resilience scores (RS) in CB; as well as the binary classification of resilience under each of the three methods. After estimating all the measures with a common dataset, we proceed to carry out a set of comparative analyses between the measures. In this section, we introduce these comparative approaches that comprise our empirical strategy.

3.1 Non-parametric distributions of the resilience capacity measures

A comparison of the non-parametric densities of the estimated resilience capacity measures' distributions under each method partially helps to determine how each method classifies the sample in terms of resilience capacity. Once generated with the same data, we want to determine how similar the distributions of the resilience capacity measures are. Do they merely differ by scaling but are otherwise similarly distributed? or are the distributions of the resilience capacity values given to the households in the sample completely different? To address this, we transform all the RCI/RS measures to a 0-100 scale and plot their non-parametric kernel densities against each other. The unitless RCIs are re-scaled to 0-100 by a mini-max procedure, while CB's RS is just multiplied by 100, representing a percentage conditional probability.

3.2 Ranking households by resilience capacity

A key comparative question to address is if the same households get consistently high or low resilience capacity values in each one of the methods. Do the methods order households similarly, even if the cardinal RCI/RS measures differ significantly? To answer this question, we use the Spearman rank correlation coefficient, a non-parametric measure of the strength and direction of the monotonic relationship between two ranked variables. With the Spearman rank correlation coefficient we can determine how likely is that the ranking of the households in one measure matches the one of the other two.

Additionally, for targeting purposes, we might be especially interested in the ranking of the households that are worst off; the ones that have the lowest estimated resilience capacity values. Toward that end, we identify the households that fall below the 20th percentile in terms of estimated resilience capacity by each method and see what percentage of those households are also under the 20th percentile in the other two. We also determine the percentage of those households that are under the 20th percentile in terms of standard well-being measures, to see how related resilience capacity measures and standard well-being outcomes are. Are we targeting the poorest when using estimated resilience capacity?

3.3 Internal consistency

Since for both RIMA and TANGO the classification of households as resilient or not is not directly related with the estimation of the resilience capacity measure, we want to determine how those two empirical identifications of resilience relate to each other. This checks the internal consistency of the resilience measures, whether they estimate resilience capacities in a way consistent with classifying households as resilient or not, since a higher RCI should presumably imply higher likelihood of being resilient. We address this with t-tests of comparisons of means between those classified as resilient and non-resilient, to see if they do have significantly different estimate RCI values. If the latent resilience capacity is correctly identified under the measure, and the binary classification of resilience accurately captures how that capacity is realized stochastically, then we expect the values between groups to differ significantly. In the case of CB, the binary classification of resilience is tied to the resilience measure itself, so we will test to caution if this is the case and both groups have significantly different probabilities of being above the specified well-being threshold.

3.4 Out-of-sample well-being prediction

For RIMA and TANGO, resilience is the capacity that ensures that shocks and stressors do not have long-lasting development consequences. For CB, resilience is the capacity that allows the unit of analysis to avoid poverty time, even in the face of shocks and stressors. All three measures are thus related, at least conceptually, to conventional well-being measures. Development goals and poverty status are usually measured in relation to normative well-being thresholds. Consequently, a natural question is whether resilient households indeed meet or exceed normative levels of well-being through time. To address this, we estimate the probability of being above or below normative levels of well-being in two successive time periods conditional on being classified as resilient or not in the first one. For example, what is the probability of being classified as resilient and being under normative levels of well-being in both time periods? Do the households classified as resilient typically fall above well-being thresholds?

Additionally, we also want to determine how close are resilience measures to current and future well-being. We address that by testing the cross-sectional and panel out-of-sample well-being predictive accuracy of each method. The out-of-sample prediction of well-being, both cross-sectional and panel, is done in two different ways. The first one is a direct relation of predicted and measured well-being outcomes. For each method, we identify the specification proposed in the original papers that relates resilience capacity and well-being, to determine which one best predicts observed well-being. In the RIMA case that will be the SEM-MIMIC model (equations 1 and 2); in TANGO, the cross-sectional OLS regression that relates well-being, the RCI and shocks (equation 4); and in CB, the main specification that yields the mean of the well-being probability density function, that relates the characteristics that build resilience capacity, shocks and former well-being to present well-being (equation 8). After estimating those models, we carry out the following forecast evaluation regression, for both the cross-sectional and panel approaches, regressing actual observed well-being for households not in the estimation sample on the well-being value predicted for those households based on the the estimation sample. Mathematically, for household i at time t,

$$Well - being_{it;real} = \delta + \rho \; Well - being_{it;predicted} + v_{it} \tag{11}$$

We then formally test the joint null hypothesis that the intercept and slope equal zero and one, respectively.

The second analysis focuses on the cross-sectional and future well-being predictive accuracy of the resilience capacity measures. For that, we generate those RCI/RS measures and estimate the following regression,

$$Well - being_{it;real} = \phi + \theta \; Resilience\ Capacity\ Measure_{it} + \omega_{it} \tag{12}$$

Since resilience capacity and well-being are not in the same units, we cannot interpret the slope and intercept as in equation 11, and so we base our comparison on the root squared mean error (RMSE) as a measure of model fit. A lower RMSE indicates a better predictor of well-being. With this exercise we want to determine how closely estimated resilience capacity tracts cross-sectional and future well-being.

The out-of-sample estimation procedure is the following. In the cross-sectional case, we estimate the resilience capacity measures and well-being outcomes with a 75% random sample of the last survey wave, maintaining the clustered structure of the survey, and using the results to predict resilience measures or well-being on the other 25% sub-sample. Then we compare that predicted resilience capacity measure or well-being, depending on the approach, to the real real measured well-being of that same 25% sub-sample. In the panel case, we estimate the resilience measures or well-being outcomes with all-but-final survey rounds and predict the resilience capacity measures or well-being outcomes in the last survey round. Then, we compare those predicted values with the real well-being measures of that last survey round, estimating equation 11 or 12 and, in the case of equation 11, testing the joint null hypothesis described there

In the panel case we go further and also run a regression of lagged and future well-being with the second-last and last survey waves. The final purpose is to see if the resilience capacity measures can be a more accurate prediction of well-being than standard lagged well-being measures.

3.5 Resilience capacity mediating the effects of shocks on well-being

In principle, according to the theoretical definition of resilience of both methods, resilience capacity mediates the negative effects of shocks on well-being. Our final comparative analysis explores if differences in the resilience measures translate into differences in that supposed mediating capacity. We address that question by estimating the following regression model,

$$Well-being_{it} = \xi + \lambda_1 SE_{it} + \lambda_2 \, RC_{it-1} + \lambda_3 SE_{it} * RC_{it-1} + \lambda_4 X_{it} + \nu_{it} \tag{13}$$

where the well-being of household i at time t is explained by a measure of shock exposure, SE_{it}, between t-1 and t, the household's resilience capacity in t-1, RC_{it-1}, and a vector of household characteristics, X_{it}. The coefficient of interest in this regression is λ_3. A positive and significant coefficient estimate will indicate that resilience capacity does mitigate the negative effects of shocks on well-being. The weakness of this test, of course, is that one cannot know the true value of λ_3 in the population so we have no benchmark against which to compare. We just have to compare against an unconfirmed belief.

CHAPTER 4

DATA SUMMARY STATISTICS AND GENERATION OF VARIABLES

4.1 The Ethiopian LSMS-ISA survey

For our comparative analysis, we use LSMS-ISA panel data from Ethiopia. Developed by the World Bank in partnership with local national statistical agencies, the Living Standards Measurement Study (LSMS) is a project designed to generate nationally representative, geographically and inter-temporally comparable measures of poverty and other well-being outcomes in developing countries. The LSMS- Integrated Surveys in Agriculture (ISA) is an extension of the LSMS project, with a stronger focus on agriculture. The surveys gather information at the household and community levels to further the understanding of the links between agriculture, other non-agricultural income generating activities and socio-economic status. To date, they have been implemented in a total of eight Sub-Saharan African (SSA) countries (World Bank, 2019_a).

The LSMS-ISA project started in Ethiopia in 2011-2012. It was originally designed to be representative of the country's rural and small-town areas [1], drawn as a sub-sample of the more comprehensive sample frame used by the Ethiopian government's Agricultural Sample Survey (ASS). The Ethiopian Socioeconomic Survey (ESS) -the name that the LSMS-ISA project receives in Ethiopia- follows a two-stage randomization procedure, with the first level at the Enumeration Area (EA), and the second at the household level (World Bank, 2019_b). A total of 333 EAs (290 rural and 43 small town) were selected in the first wave of the panel survey. With 12 households for EA, the total sample size of the first survey wave was 3,969 households. At the geographical level, the ESS is representative of the four most populated Ethiopian regions: Amhara, Oromiya,

[1] A small town corresponds to towns with a population of less than 10,000 people, based on the 2007 Population Census

Southern Nations, Nationalities and Peoples (SNNP) and Tigray. The sample is not representative of each one of the other smaller regions, Afar, Benshangul Gumuz, Dire Dawa, Gambella, Harari, and Somalie individually, but it can be considered representative of them as a whole if considered in one ”smaller regions” category.

A total of three waves comprise the final panel dataset. The second wave was collected in 2013-2014 and the third and, to date, last one in 2015-2016. From the original sample of 3,969 households, a total of 3,776 were re-interviewed in the second wave, and 3,699 in the third, which corresponds to attrition levels of 5 and 7%, respectively. The total sample size of those last two waves was increased on 1,486 households to incorporate urban areas, but we do not include that additional sample in our analysis because it is not available for all waves. For each wave, the information is collected in three visits, designed to follow the Ethiopian agricultural calendar. The first visit takes place in the post-planting period (Sept-Oct), the second in the middle of the agricultural season (Nov-Dec) and the third in the post-harvest period (Jan-Mar in first wave and Feb-Apr in the other two). In the first visit, the post-planting agricultural questionnaire is collected. In the second, the household completes a livestock questionnaire. And in the last one, the post-harvest agricultural questionnaire and the household and community questionnaires are completed. Therefore, all the time-variant information on household and community characteristics is mostly representative of the post-harvest period. A summary of some key facts of each survey wave can be found in Appendix C.

4.1.1 Outcome variables

The analysis of resilience always needs to relate to some well-being outcomes, since reaching adequate and stable levels of well-being is the ultimate goal. We use two different measures of well-being; total per adult equivalent consumption expenditures, to reflect poverty status, and the

Food Consumption Score, directly related to food security.

For the consumption expenditures, we use real per adult equivalent annualized total expenditures (RAEC), which are available at the household level for each one of the Ethiopian LSMS-ISA waves. The consumption aggregates are calculated from the household's food consumption on 25 different food items; their expenditures on 11 semi-durable goods (e.g., batteries or soap) and 12 durable goods (e.g., clothes). The recall time for each one of the categories is: seven days for food items, one month for the expenditures in semi-durable goods and a year for the durable goods. To allow for the spatial comparability of the prices, the survey incorporates a Fisher price index, using price data from both household purchases and the market price survey. For each household, food consumption expenditures are estimated by multiplying the total amount of each food consumed by the median price paid for that item at the EA. The median price is obtained from the own survey, corresponding to the median of the self-reported price payed for that item across all the households in the EA. Apart from purchased food, the consumption aggregates also include the value of the food that was obtained by own production and through gifts. Non-food expenditures include the self-reported purchased value of the semi-durable and durable goods, along with some education expenditures (e.g., amount spent on school fees, school books, uniforms, etc.).

The real per adult equivalent consumption expenditures are deflated, adjusted for temporal variation in prices using the Consumer Price Index (CPI), constructed from information on monthly inflation from the information in the online platform of "Trading Economics" (TradingEconomics, 2019). They are also adjusted for spacial variation using a Fisher food price index constructed from the price data available in the datasets. The real adjustments for Ethiopia thus bring everything across the three waves back to Feb 2012 rural/small town prices. The number of equivalent adults in the household is estimated from the age and the gender of each household member, using the equivalences outlined in Table 1.

Table 1: Conversion Factors for Household Equivalency

Age Range	**Male**	**Female**
0	0.30	0.30
1	0.46	0.46
2	0.54	0.54
3-4	0.62	0.62
5-6	0.74	0.70
7-9	0.84	0.72
10-11	0.88	0.78
12-13	0.96	0.84
14-15	1.06	0.86
16-17	1.14	0.86
18-29	1.04	0.80
30-59	1.00	0.82
60 or more	0.84	0.74

Source: Dercon & Krishnan, 1998

The FCS was originally developed by the World Food Program (WFP), as a population-level measure of food security that captured not just total caloric intake, but also diet quality with the measurement of diet diversity. Nevertheless, the measure has been validated in terms of its relationship with total caloric intake, but not in terms of total nutrient acquisition, so its value for assessing diet quality is yet to be proved (Leroy et al., 2015; Weismann et al., 2009). The FCS is estimated with seven-day recall data in the consumption of different food items. It records the additive daily frequency of the consumption eight food groups. A total of one unit per day can be added in each food group for the daily consumption of any food item classified in that group. This frequency consumption additive units are then weighted by a coefficient specific for each group [2]. The weighted scores are added to create the final numeric index (WFP, 2007).

[2] The food groups (and weights) are as follows: staples (2), pulses (3), vegetables (1), fruit (1), meat and fish (4), milk (4), sugar (0.5), oil (0.5).

4.1.2 Well-being normative thresholds

To determine adequate levels of well-being from the consumption aggregates we use a consumption poverty line generated to match the design of the survey. The total number of items included in each one of the expenditure categories is limited, and therefore an externally generated, more comprehensive poverty line such as the World Bank global poverty line or those calculated by national governments can lead to an inaccurate classification of the population in poor/non-poor categories due to non-comparability. The total consumption poverty line is estimated in two steps, first generating a poverty line for the food expenditures, and then using that as a reference to generate the non-food poverty line, and finally combining both of them. For food consumption, the poverty line is estimated using a food basket of 2,700 kcal/day. The products included in the basket vary slightly by wave and reflect the diet of the poor and middle-income households (bottom 70% of the real per adult equivalent food consumption distribution). The non-food expenditure poverty line uses the food consumption poverty line as a base. Specifically, the Ravallion method is used (Ravallion, 1992), which averages real non-food expenditure for those households whose food consumption is within x% of the food poverty line, for x ranging between one and ten. That average of real non-food consumption expenditures is added to the food consumption poverty line to create the final total poverty line. For the FCS, we use the generally accepted score of 35 as a threshold that determines adequate food intake (Wiesmann et al., 2009).

Figure 1 shows the non-parametric kernel densities of the FCS (left) and consumption aggregates (right). Table 2 shows, by wave, the summary statistics of the two measures of well-being used in this paper -the real per adult equivalent consumption expenditures and the FCS- jointly with the well-being thresholds and the percentage of households in the sample below those thresholds. Both consumption expenditures and poverty lines are expressed in constant prices (February 2012 prices), so the comparison between waves is straight forward.

Finally, it is noticeable that the internally estimated poverty line, built to match the design of the survey, is somewhat odd in relation to the Ethiopian sociopolitical context of the years where the survey took place. The estimated poverty line decreases over time; the one of waves two and three being almost equal to one another and both lower than the one of the first wave. However, that time period is reported to be of substantial, widespread, economic growth and total poverty reduction in Ethiopia (United Nations, 2018). We have not going further in addressing this issue, to investigate if that reported economic growth was enough widespread to reach the rural communities represented in the data used in this thesis, but, nevertheless, this discrepancy does not represent a problem for our comparative analysis, since the poverty line used to generate the qualitative results of this thesis is common for all the three measures.

Table 2: Summary statistics of well-being outcome measures. Normative well-being thresholds

	Wave 1	**Wave 2**	**Wave 3**
RAEC	4,612 (2,741)	3,916 (2,373)	3,302 (2,138)
Poverty line (birr.)	3,973	3,478	3,491
Poor	48%	51%	62%
FCS	41 (17)	42 (17)	42 (17)
FCS limit score	35	35	35
Food insecure	33%	32%	32%
Observations	3,969	3,776	3,699

Note: Standard deviation in parenthesis

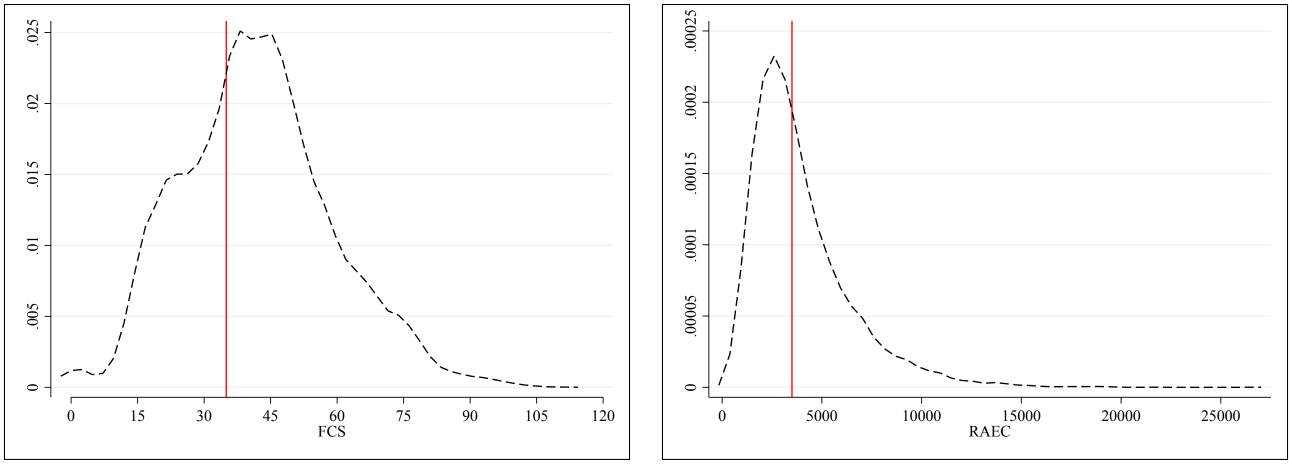

Figure 1: Non-parametric densities of the FCS (left) and RAEC (right), all waves

4.1.3 Generation of resilience-related variables

One important aspect of any empirical resilience model is the criteria used behind the selection of the variables that are thought to build resilience capacity. Resilience capacity is an intrinsically complex, multidimensional concept, operationalized by a set of characteristics at different aggregation levels (household, community). Informed by previous research in the areas of resilience, vulnerability and food security, both RIMA and TANGO propose a conceptual guideline on what set of variables should be included in the estimation of the resilience capacity index (RCI). Cissé and Barrett (2018) leave this choice to the researcher's criteria, but advocate for the contextualization of the selection to the particular conditions where the data was collected.

Table 3 shows the variable selection in the original RIMA, TANGO and CB papers (Smith and Frankenberger, 2018; Cissé and Barrett, 2018; FAO, 2016). The rows in bold text represent the final factor analysis indexes generated with the variables included in each category. It is important to note that many of the variables in TANGO are themselves additive or factor analysis indexes. So these capacities represent indexes of indexes. From the three methods, TANGO is the one that needs the richest dataset for the estimation of the RCI. The concepts captured in RIMA are already contained in TANGO's variable selection.

Table 3: Resilience capacity variable selection in the original papers

TANGO	RIMA-II	CB
1. Absorptive Capacity	**1- Access to Basic Services (ABS)**	Left to the researchers' criteria
Bonding Social Capital Asset Ownership Cash Savings Access to informal safety nets Support for disaster risk reduction	Infrastructure index Distance to vet. services Distance to primary school Distance to health services Distance to product market Distance to non ag. market	
2. Adaptive Capacity	**2- Assets (AST)**	
Bridging Social Capital Linking Social Capital Aspirations and Confidence to Adapt Livelihood Diversity	Agricultural Assets Wealth Index Tropical Livestock Units(TLU)	
Human Capital Access to Information	**3- Social Safety Nets (SSN)** Transfers	
3. Transformative Capacity	Other Transfers Scholarships	
Bridging Social Capital Linking Social Capital	**4- Adaptive Capacity (AC)**	
Governance (Quality) Access to markets Access to services Women's Empowerment	Participation Index HH years of education Dependency ratio Crop Diversification Index	

NOTE: in bold, factor indexes built from the group of variables under each one of them

A crucial part of our empirical strategy relies on the comparability of the results between models, so that the differences between them, if any, can be attributed entirely to their intrinsic construction. That is the main reason behind using a common dataset for the estimation of the three methods. That common dataset should be maintained through all the process, starting from the generation of a comprehensive set of variables from which all the models can be estimated, so that any differences do not arise from differences in the data or variables used. Following as closely as possible the original theoretical criteria and methodological procedure outlined in each one of the methods for the generation of resilience capacity building variables, we make an effort to adapt those guidelines to the LSMS-ISA data. Since the TANGO method is the most demanding of the three in terms of data requirements, we base our adaptation procedure to match their requirements. It is easy then to pick the necessary variables from that extensive dataset to generate the other models according to their original guidelines.

We present in Table 4 the summary statistics of the variables from the LSMS-ISA dataset used for the estimation of the resilience capacity building variables of the three models. We classify them following the TANGO method and FSIN's guidelines, depending on if they form part of the absorptive, adaptive or transformative capacities. The fifth column specifies in which model or models that variable is included in the direct estimation of the resilience capacity measure, either as the variable itself or combined with other variables in a more complex index. A detailed explanation on how those variables are grouped to form the underlying variables and the three resilience capacity indexes in TANGO and four pillars in RIMA can be found in Appendix D. Recall that TANGO uses an altered, unorthodox version of the regular principal component factor analysis to generate their variables (see Appendix B), that we follow when generating factor indexes for computing the TANGO measure. For RIMA, we just use regular principal factor analysis, as they do in their original paper. Likewise, for indexes included in CB we use standard factor analysis methods.

Some steps of our empirical estimation also require the use of control variables or household characteristics. Since, again, the TANGO method is the one using a more comprehensive set of control variables, we use their selection in our models. The summary statistics of those control variables are also presented in Table 4.

Table 4: Summary statistics of Ethiopian LSMS-ISA selected variables

	Mean	SD	Min	Max	Model
Control Variables					
Wealth Quintile	2.96	1.51	1	5	CB
Age of hh head	45.71	15.49	15	100	CB
Sex of hh head (1=male, 0=female)	0.74	0.44	0	1	CB
Primary education, completed (hh head)	0.22	0.41	0	1	CB
Secondary education, completed (hh head)	0.09	0.29	0	1	CB
College education, completed (hh head)	0.05	0.21	0	1	CB
Household Size	4.93	2.38	1	16	CB
% of males in hh (age 0-4)	0.07	0.11	0	0.67	CB
% of males in hh (age 5-15)	0.15	0.16	0	0.80	CB
% of males in hh (age 16-65)	0.24	0.20	0	1	CB
% of males in hh (age ¿66)	0.03	0.10	0	1	CB
% of females in hh (age 0-4)	0.06	0.11	0	0.6	CB
% of females in hh (age 5-15)	0.14	0.16	0	1	CB
% of females in hh (age 16-65)	0.28	0.20	0	1	CB
% of females in hh (age ¿66)	0.03	0.15	0	1	CB
Absorptive Capacity					
Prob. of borrowing from close circle (EA level)	0.15	0.18	0	1	T, CB
% of hh receive informal transfers (EA level)	0.17	0.19	0	1	T, CB, R
Value informal transfers (hh level)	138. 66	951.96	0	59371.87	T, CB, R
Durable Asset Index	0.00	2.09	-1.43	20.58	T, CB, R
Productive Asset Index	0.00	1.29	-2.60	15.05	T, CB, R
Livestock holdings (TLU)	4.16	17.51	0	742.1	T, CB, R
Access to agricultural land	0.80	0.40	0	1	T, CB, R
% of hh receive any type of assistance (EA level)	0.19	0.30	0	1	T, CB, R
Members in community migrate seasonally	0.73	0.44	0	1	T, CB
Distance to health center (EA level)	1.02	4.08	0	40	T, CB
Adaptive Capacity					
Presence of cooperative in the EA	0.18	0.39	0	1	T, CB, R
Share of males participate in the coop.	3.19	12.23	0	100	T, CB
Share of females participate in the coop.	1.74	6.65	0	75	T, CB
Family member works in government/political party	0.06	0.24	0	1	T, CB
Num. of distinct income sources (out of 7)	0.13	0.37	0	3	T, CB, R
Max. num. of years of education completed (all hh members)	4.92	4.06	0	18	T, CB, R
Num. of members with disability (hh)	0.36	0.73	0	12	T, CB, R
Num. of literate members (hh)	1.97	1.79	0	12	T, CB, R

	Mean	SD	Min	Max	Model
Access to information	0.79	0.94	0	4	T, CB, R
Transformative Capacity					
Distance to main periodical market (Km, EA level)	7.66	16.99	0	250	T, CB, R
Distance to primary school (Km, EA level)	1.17	10.41	0	300	T, CB, R
Distance to secondary school (Km, EA level)	12.99	24.52	0	500	T, CB, R
Distance to pharmacy (Km, EA level)	6.61	19.96	0	500	T, CB, R
Distance to bus (Km, EA level)	17.61	30.81	0	568	T, CB, R
Distance to paved road (Km, EA level)	43.12	58.83	0	675	T, CB, R
Distance to extension agent (Km, EA level)	1.04	7.95	0	200	T, CB, R
% of hh (EA level) with enterprises where its owned by women	0.39	0.38	0	1	T, CB
% of hh (EA level) with loans where women decides over loan	0.21	0.30	0	1	T, CB
% of hh (EA level) that receive income where women decides over income	0.34	0.34	0	1	T, CB
% of hh (EA level) that own ag. land where women owns all/part of the land	0.29	0.16	0	1	T, CB
Needs for which community (EA) ask governmental institutions	5.42	2.80	0	11	T, CB
Needs for which governance consults community (EA)	5.56	2.78	0	11	T, CB
Mean level of need addressed (EA level) (1=none, 5=addressed)	3.59	1.19	1	5	T, CB

NOTE: T=TANGO; CB= Cissé and Barrett; R=RIMA-II

The variables here are just listed once, but some can appear in multiple categories. For a detailed explanation consult Appendix D

Finally, we also present the summary statistics of the final indexes for the RIMA (Table 5), TANGO and CB methods (Table 6), generated by the different data reduction procedures (factor analysis and additive indexes) prescribed by TANGO and RIMA with the variables from Table 4 (to see the selection of variables used to estimate each one of them refer to Appendix D). Since CB leaves the variable selection to the researcher's criteria, for the sake of comparability, we include in their model the same set of resilience capacity building and control variables used in the estimation of TANGO's method. Table 6 shows the summary statistics of the indexes generated with the TANGO factor analysis method and regular principal component factor analysis. All the factor analysis indexes are rescaled from 0 to 100 to facilitate their interpretation. The factor loadings of all the factor analysis indexes can be found in Appendix E.

Table 5: Summary statistics of resilience capacity variables used in RIMA (pillars)

Access to Basic Services (ABS)	21.87 (21.63)
Assets (AST)	34.53 (10.58)
Social Safety Nets (SSN)	4.22 (4.43)
Adaptive Capacity (AC))	29.94 (18.28)
NOTE: Standard deviations in parenthesis. The 4 pillars are factor indexes (0-100 range).	

4.2 External measures of drought

Resilience is only understood in the face of shocks and stressors. In this paper, we focus on drought as the main event potentially causing negative effects in household well-being. Although self-reported information on shock exposure is available in the LSMS-ISA datasets, we choose to use an external measure of exposure to drought, the normalized difference vegetation index (NDVI). With this, we try to avoid the possibility of biasing our results by introducing non-classical measurement error via the self-reported data. Non-classical measurement error is a type of measurement error that, because it is related to some of the variables included in the model, can bias estimation results (Abay et al., 2019).

The NDVI is a satellite remote sensing measure that captures the amount of photosynthetically active radiation that the vegetable cover absorbs. The index ranges between -1 and 1, with greater values indicating higher densities of green leaf cover. The NDVI is therefore a measure of plant stress and overall growth. To relate it to abnormal weather events, such as drought, the common procedure is to generate a normalized NDVI index, that measures positive and negative standard deviations from a mean value of NDVI averaged with data from previous years. We add to our LSMS-ISA data geographically referenced measures of the NDVI at the EA level that correspond to the average value of the measure for months prior to the post-harvest period. We then generate the standardized value of the NDVI for those critical months of the crop growing season, with

Table 6: Summary statistics of resilience capacity variables used in TANGO and CB

	TANGO FA	Regular FA
Absorptive Capacity	4.79 (4.49)	4.43 (3.81)
Bonding social capital (0-100)	4.01 (3.89)	3.83 (3.71)
Asset Index (0-100)	7.15 (3.57)	34.54 (10.58)
Access to informal SN (0-100)	5.68 (5.83)	4.21 (4.43)
Disaster mitigation (0-100)	60.67 (27.23)	80.50 (27.47)
Adaptive Capacity	21.15 (16.85)	23.58 (15.14)
Bridging social capital (0-100)	7.15 (17.74)	5.94 (15.49)
Linking social capital (0-1)	0.06 (0.25)	0.06 (0.25)
Livelihood Diversity (0-7)	0.14 (0.37)	0.14 (0.37)
Human Capital (0-100)	24.74 (17.06)	24.51 (17.22)
Access to Information (0-4)	0.79 (0.94)	0.79 (0.94)
Transformative Capacity	27.04 (12.81)	22.03 (12.63)
Distance to per. market	7.66 (16.99)	7.66 (16.99)
Access to services (0-7)	3.53 (1.17)	3.53 (1.17)
Women′s empowerment (0-100)	38.64 (18.72)	32.86 (15.44)
Quality of governance (0-100)	53.10 (20.90)	52.08 (22.27)

NOTE: Standard deviations in parenthesis. Some of the variables form part of multiple categories in the TANGO method (see table 2)

the average values from all the prior years available in the data. We therefore try to relate abnormal negative deviations in the average NDVI with negative effects in average crop yields at the EA level. Although with the construction of that variable we imply this relationship, we do not formally test it in our sample (e.g., do not compare average annual EA crop yields with different NDVI scores).

To generate a more precise measure of drought, we create a new continuous variable, based on the standardized NDVI, but truncated at values of the measure that likely do not reflect a serious, adverse shock. More specifically, we transform to zero all the values higher than 1.5 standard deviations below the mean. With that, we want to capture better the anomalous deviations at the low end of the distribution that can reflect negative shocks in crop yields due to drought. This continuous truncated measure of drought is the one that we include in the estimation of the models. Figure 2 shows the distribution of the mean standardized NDVI in our sample. The red line indicates the value of -1.5 standard deviations from which we truncate the measure to capture the effects of drought.

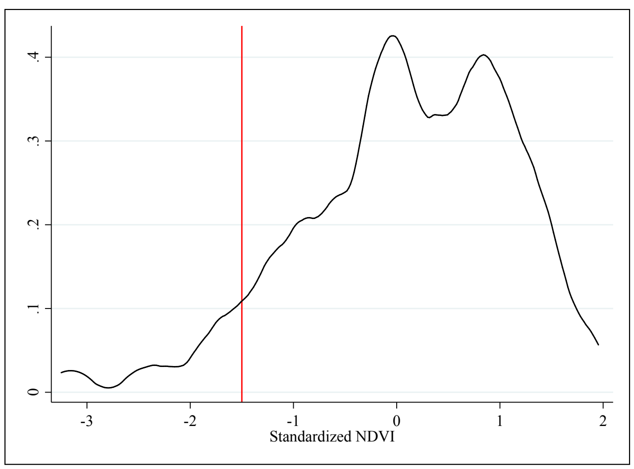

Figure 2: Non-parametric density of the mean standardized NDVI, all waves

CHAPTER 5

RESULTS

5.1 Resilience capacity measures

Since our comparative analysis is mainly focused on the three resilience capacity measures, we present here the key results of their estimation. How we present those depends on their original estimation procedure.

For RIMA, we present in Table 7 the results of the SEM-MIMIC model that is used for the estimation of the RCI. The four pillars and well-being outcomes were normalized before the estimation of the model, so the results are expressed in units of standard deviation units. Standard errors are clustered at the EA level and we use sample weights.

Conversely to what was found in previous RIMA analyses, where all the pillars had a significant effect on the RCI (FAO, 2016; d'Errico et al. 2018), in our sample, the only pillar statistically significant in determining the resilience capacity measure is Adaptive Capacity. This pillar is formed by variables related to human capital, livelihood diversity and access to information. Comparing to these other studies, it is important to note that in those the error terms were not clustered according to the complex structure of the survey, even though d'Errico et al., (2018) use also LSMS-ISA data when estimating RIMA's RCI. Failing to cluster the standard errors when those are correlated within cluster can lead to smaller standard errors that what should be, and therefore lower p-values and a higher likelihood of having statistically significant coefficients (Cameron and Miller, 2015). Due to the design of the LSMS-ISA Ethiopian survey, with a relatively large number of clusters (originally 333 EAs) and few observations per cluster (in average, 12 households per EA), we consider the clustering of the standard errors appropriate, and therefore we do so when estimating the

RCI. The effect of the RCI on the FCS is positive and significant while the coefficient of the other measure of well-being, the RAEC, had to be fixed to one to allow for the estimation of the model.

We present two goodness-of-fit statistics of the SEM-MIMIC model, the standardized root mean squared residual (SRMR) and the coefficient of determination (CD). Note that because we cluster the standard errors, we are not able to estimate other goodness-of-fit tests that RIMA presents in their original paper, such as the root mean square error of approximation (RMSEA), the comparative fit index (CFI) or the Tucker Lewis index (TLI). The SMRS indicates a good model fit, close to zero and below 0.08 (Hu & Bentler, 1999). The coefficient of determination can be interpreted as an R-squared.

Table 7: Multiple Indicators Multiple Causes (MIMIC) model, RIMA

Structural Component	
Access to Basic Services (ABS)	0.009
	(0.029)
Assets (AST)	0.008
	(0.041)
Social Safety Nets (SSN)	0.034
	(0.034)
Adaptive Capacity (AC)	0.215***
	(0.032)
Measurement Component	
Real annualized consumption expenditure (RAEC)	1
	(0.000)
Food Consumption Score (FCS)	1.267***
	(0.231)
Goodness-of-fit statistics	
SRMR	0.024
CD	0.147
Observations	10,313

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

For the TANGO method, table 8 shows the factor loadings of the RCI estimation by factor analysis (the factor loadings of the variables that were included to generate the absorptive, adaptive and transformative capacity indexes can be seen in Appendix E). Recall that TANGO uses a modified unorthodox version of the principal factor analysis, so those factor loadings are just estimated with

Table 8: Factor loadings of the three capacities that form the RCI in TANGO

	Factor Loading
Absorptive Capacity	0.252
Adaptive Capacity	0.534
Transformative Capacity	0.562

NOTE: Refer to Appendix B for an explanation of TANGO's FA method.
Refer to Appendix D and E for the variables that form each capacity and their factor loadings.

data from the first wave (refer to Appendix B for a detailed explanation of TANGO's factor analysis method). From the three components of the index, absorptive capacity is the one less correlated with the factor, while the adaptive and transformative capacity factor loadings are higher and more similar with each other (see Appendix D for the variables that form the absorptive, adaptive and transformative capacity factor indexes).

For the CB method, the resilience score is derived from the estimation of equations (8) and (9) (pg.12). We present in Table 9 the results of those two regressions that yield the conditional mean and conditional variance of the well-being p.d.f from which the resilience score is estimated. For simplicity, the results of the regressions are presented with just the resilience related variables included in the estimation, and no the control variables. A complete table with all the variables included in the model can be seen in Appendix F. We run a total of four regressions, two for each well-being outcome of interest. For the sake of comparability, and because CB leaves the selection of the control and resilience capacity building variables to the researcher's criteria, we use the same set of variables as TANGO, but with two differences. First, we include the variables in individual form; we do not group them in absorptive, adaptive and transformative capacity indexes. And second, we use regular principal factor analysis to estimate the variables that are factor indexes, rather than using TANGO's nonstandard factor analysis procedure. Since resilience is just understood in the face of shocks, we also include in model the measure of exposure to drought. We use the predicted values of those two regressions to generate the conditional mean and the conditional variance that corresponds to the well-being conditional p.d.f, unique for each household. Like in CB's original paper, we assume a gamma probability density function.

Finally, we present in Table 10 the summary statistics of the three resilience measures. For comparison purposes all are converted to a 0-100 scale; the TANGO and RIMA measures using a mini-max procedure, as outlined in the original papers, and CB multiplying the conditional probability by a hundred, obtaining conditional probability percentages. The interpretation of these results should be done carefully, since all the measures do not represent the same concepts. For RIMA and TANGO, the comparison is more straightforward since, although with differences in variable selection and estimation procedure, both are unitless indexes generated as latent variables via data reduction techniques (factor analysis). Both indexes are indeed fairly similar, centered around values of twenty, and their means just differing by three points.

The comparison of those two measures with CB's is not as straightforward. Under CB, the resilience capacity measure, the resilience score, is not an unitless index but a predicted conditional probability of being above a pre-defined well-being threshold. The interpretation of those resilience measures should therefore take in account the normative thresholds from which they were generated, which in our case are the poverty line for the RAEC and the limiting score of 35 for the FCS. In our sample, according to the CB method, the average probability of being non-poor is 42% and of being food secure -according to the FCS- is 62%. Consequently, differences between CB's and TANGO and RIMA resilience measures are somehow expected, since they represent different ways of identifying resilience capacity empirically. Even though in all the cases the higher the value the better, the numbers resultant of the estimation of CB's measure are not directly comparable with the ones of the RIMA and TANGO indexes.

Table 9: Results of regressions to generate mean (left) and variance (right) of well-being p.d.f

	RAEC	**FCS**
RAEC, (t-1)	0.40***	
	(0.10)	
RAEC, (t-1), sq. (*10^4)	-0.05	
	(0.17)	
RAEC, (t-1), cub. (*10^8)	-0.07	
	(0.08)	
FCS, (t-1)		0.30*
		(0.18)
FCS, (t-1), sq. (*10^2)		-0.11
		(0.42)
FCS, (t-1), cub (*10^3)		0.01
		(0.03)
Drought	61.86	-0.25
	(68.08)	(1.19)
Bonding capital	-27.81	1.26
	(150.60)	(1.37)
Bridging capital	-50.84	-0.12
	(58.99)	(0.71)
Linking capital	465.6*	3.67*
	(276.50)	(2.16)
Asset Index	182.90	1.93**
	(115.50)	(0.95)
Livelihood Diversity	15.07	-1.859*
	(86.99)	(0.95)
Access to services	20.19	0.411
	(55.13)	(0.53)
Human capital	-40.22	0.682
	(63.38)	(0.806)
Access to information	362.5***	1.508***
	(54.00)	(0.49)
Women's empowerment	-141.40	0.05
	(120.60)	(1.34)
Quality of governance	-66.01	0.27
	(63.88)	(0.65)
Informal Safety Nets	168.70	2.84**
	(161.60)	(1.42)
Disaster risk reduction	-977.20	-27.56***
	(749.40)	(9.03)
Constant	1,874***	23.93***
	(336.90)	(3.19)
Controls	Yes	Yes
Observations	6,137	6,460
R-squared	0.29	0.21

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

	res.(RAEC)	**res.(FCS)**
RAEC, (t-1) (*10^3)	0.15	
	(0.09)	
RAEC, (t-1), sq. (*10^8)	-0.06	
	(1.27)	
RAEC, (t-1), cub. (**10^7)	-0.01	
	(0.01)	
FCS, (t-1)		0.01
		(0.01)
FCS, (t-1), sq. (*10^3)		0.27
		(0.26)
FCS, (t-1), cub. (*10^4)		0.03
		(0.03)
Drought	-0.06	-0.04
	(0.08)	(0.05)
Bonding capital	-0.11	0.07
	(0.13)	(0.10)
Bridging capital	-0.03	0.05
	(0.04)	(0.04)
Linking capital	-0.34	0.05
	(0.33)	(0.09)
Asset index	-0.03	-0.15**
	(0.09)	(0.06)
Livelihood diversity	-0.06	-0.08
	(0.11)	(0.07)
Access to services	-0.05	-0.02
	(0.04)	(0.03)
Human capital	-0.12	-0.05
	(0.08)	(0.05)
Access to information	0.12**	0.06**
	(0.05)	(0.03)
Women's empowerment	-0.05	0.13**
	(0.08)	(0.06)
Quality of governance	-0.04	-0.02
	(0.05)	(0.04)
Informal safety nets	0.27*	-0.19**
	(0.14)	(0.10)
Disaster risk reduction	0.64	0.19
	(0.55)	(0.39)
Constant	14.53***	5.63***
	(0.37)	(0.19)
Controls	Yes	Yes
Observations	6,137	6,460

Robust standard errors in parentheses. res.= residuals
*** p<0.01, ** p<0.05, * p<0.1

Table 10: Summary statistics of the three resilience measures

		Mean	SD	Min	Max
TANGO	Resilience Capacity Index (RCI)	19.84	9.51	0	100
RIMA	Resilience Capacity Index (RCI)	22.22	12.47	0	100
CB	Resilience Score [RAEC]	42.31	22.20	0.59	100
	Resilience Score [FCS]	64.85	17.89	9.51	99.43

5.2 Non-parametric densities of the resilience capacity measures

We start the comparison of the three resilience capacity measures by plotting their non-parametric kernel densities for the last survey wave (Figure 3). Concurring with their measures of central tendency and dispersion, RIMA and TANGO show similar distributions, centered close to the same point. As expected, the densities of the CB probability measures are very different from the other two methods, but again, CB measures do not represent exactly the same underlying concepts as RIMA and TANGO indexes. In CB, the probability of being non-poor is skewed to the lower half of the whole range, and the opposite happens for the probability of being food secure. Again, for the interpretation of these results, it should be taken in account that CB's measures represent the conditional probability of being above a pre-specified well-being threshold, and they are therefore sensitive to the well-being normative threshold selection.

5.3 Ranking of households by resilience capacity

An important empirical aspect is how similarly those measures rank households in terms of resilience capacity, and how that ranking relates to actual levels of standard well-being measures. We address this by estimating the Spearman rank correlation coefficients between the three resilience measures and real well-being outcomes (Table 11). Consistent with what we saw in the distributions and summary statistics of the three measures, RIMA and TANGOs household ranking in

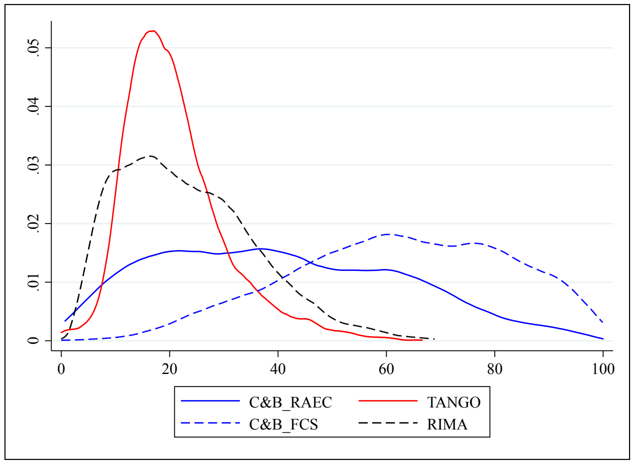

Figure 3: Non-parametric densities of the resilience capacity measures, wave 3

terms of resilience capacity is strongly correlated (corr.=0.70). Also consistent with the differences that our results previously showed, CB's resilience measures are more dissociated from the other two methods, with correlations always under 0.50 for both FCS/RAEC CB resilience scores. That implies that households that are getting relatively high or low values of resilience capacity under RIMA or TANGO, do not necessarily get also a high or low value in CB.

Regarding well-being measures, CB's resilience scores show the higher rank correlations, with values of 0.55/0.41 for the RAEC/FCS resilience scores, respectively. This is in line with the internal construction of the measure as the probability of being above a normative well-being threshold (higher value of well-being, higher probability of being above the threshold). RIMA and TANGO are much more dissociated from actual well-being status, the correlations being in all cases lower than 0.25.

Summarizing, households get ranked reasonably similarly in terms of resilience capacity when using TANGO and RIMA, a classification that differs more from the CB method. CB's resilience score is more related to actual well-being, so a ranking of the households that follows CB's resilience measures will be a closer representation of actual well-being status.

Table 11: Spearman rank correlation coefficient matrix

	RIMA	TANGO	CB[FCS]	CB[RAEC]	FCS
TANGO	0.702				
CB[FCS]	0.439	0.391			
CB[RAEC]	0.427	0.308	0.349		
FCS	0.229	0.221	0.410	0.167	
RAEC	0.269	0.196	0.266	0.548	0.361

NOTE: All the coefficients are statistically significant at the 0.01 level. Data from wave 3.

For targeting purposes, we might be most interested on how the measures identify the worst off, or the households in the sample with lower values of resilient capacity. We analyze that by comparing the households in the lowest 20% of the resilience capacity distributions. For each measure, we determine the percentage of households that are in the bottom 20% of that resilience capacity measure that are also in the bottom 20% of the other two measures of resilience capacity and well-being outcomes (Table 12). Our results agree with what we saw before with the Spearman rank correlation coefficients. The highest correspondence between households in the bottom 20% regarding resilience capacity is between the TANGO and RIMA measures. 60% of the households that are in TANGO's bottom 20% are also in RIMA's bottom 20%. TANGO and RIMA classifications are closer to the CB resilience score when FCS is the outcome measure, but the percentage of shared households is always lower than 50%. CB's ranking of households by resilience score is again closer to actual well-being measures, with 45% of the households that are in the bottom 20% of the CB's RAEC resilience score also in the bottom 20% of the RAEC distribution, and 40% when FCS is the outcome variable of interest.

To conclude this section, we want to highlight two main points derived from our results that have important consequences for those using these measures empirically. First, the ranking of the households differs between resilience capacity measures, which implies that, if those are used for targeting purposes, a different group of households might be chosen under each method. And second, both RIMA and TANGO resilience capacity measures are fairly dissociated from standard well-being outcomes, while CB more closely corresponds with actual well-being status. Under TANGO and RIMA, the worst off in terms of resilience capacity are not necessarily the same group of households with lower values of standard well-being measures.

Finally, we want to briefly deepen on the already existing concern about the dissociation between some resilience capacity measures and well-being outcomes that our results just showed. Béné et al., (2012) raised the danger of resilience capacity not being a pro-poor measure, introducing the concept of ”adaptive preference”, or how people adapt to deteriorating living conditions. Our results indicate that just looking at resilience capacity measures, specially under the RIMA and TANGO methods, is not enough to determine the underlying well-being status of the analyzed population, underscoring the Bene et al., (2012) concern that focusing on resilience might divert attention from the worst-off according to standard well-being outcomes if the resilience measure is not designed carefully so as to correspond tolerably to conventional well-being measures.

Table 12: Comparison of hhs in the bottom 20% of the resilience capacity measures

	Bottom 20% in Wave 3					
	TANGO	**RIMA**	**CB[RAEC]**	**CB[FCS]**	**RAEC**	**FCS**
TANGO	100%	51%	26%	40%	23%	23%
RIMA	60%	100%	24%	42%	26%	29%
CB[RAEC]	25%	20%	100%	29%	38%	20%
CB[FCS]	40%	36%	31%	100%	28%	32%
RAEC	26%	25%	45%	32%	100%	33%
FCS	29%	32%	26%	40%	37%	100%

5.4 Internal consistency checks

Recall that in each one of the methods, there are two ways of identifying resilience empirically. One is with the estimation of the resilience capacity measure, which yields a numerical value for each household. The other is with the binary classification of households as resilient or not.

Intuitively, those two dimensions are expected to be related, with households identified as resilient having a higher resilience capacity. Nevertheless, under the RIMA and TANGO methods, that binary resilience classification is not done with the previously estimated RCI, but rather looking at the fluctuations in standard well-being measures between survey waves. For RIMA and TANGO, resilient households are those that manage to not suffer a loss in well-being between consecutive time periods after being affected by a shock. Consequently, we want to determine the internal consistency of those two dissociated ways of identifying resilience empirically, to see if households identified as resilient by the well-being outcomes have indeed higher resilience capacity values. In the CB method, the identification of resilient households is tied to the resilience score. Under CB, resilient households have a resilience score higher than a pre-specified probability cut-off. In this case, we will expect internal consistency by construction, with resilient households having a greater conditional probability or resilience score.

An important aspect for comparison purposes between the three methods is to have a similar percentage of households in the whole sample classified as resilient and non-resilient. Since, the probability cut-off of the CB method is arbitrarily decided by the analyst/practitioner we set that percentage likelihood for CB to the level that generates the same proportion of resilient and non-resilient households as by the binary classification of resilience under RIMA and TANGO. This is to say that these are not the same households, just the same percentage assigned to the resilient and non-resilient groups. In our sample, that probability threshold for the CB method is equal to 65% for the FCS and 40% for the consumption expenditures, which yields a sample where 43% of the households are considered resilient in terms of consumption expenditure and 54% in terms of FCS. This means that, under the CB method, resilient households have a 65% or above probability of being above a 35 FCS value when FCS is the well-being outcome of interest. When consumption expenditures is the well-being measure, resilient are households that have a 40% or higher probability of being above the poverty line.

We perform t-tests of comparison of means to analyze if resilient and non-resilient groups under each classification method have significantly different resilience capacity values. Table 13 presents the results of those t-test for the resilience score measures between the groups classified as resilient or not by CB's binary classification method, using data from the last survey wave. Note that the classification as resilient or not is tied to the well-being outcome of interest, and therefore there are two resilient and non-resilient groups, depending on the probability cut-off to identify resilient households. For example, resilient households in Table 13 under the FCS columns are households that have a probability of being above a FCS of 35 higher than 65%. Under the RAEC columns, resilient households are the ones whose probability of being non-poor is higher than 40%. For each group classified under a certain well-being outcome, we compare the two resilience scores, the one representing the conditional probability of being food secure (CB[FCS]) and the one of being non-poor (CB[RAEC]). However, a priori, just the resilience score that was originally estimated for that particular well-being measure has to be internally consistent. The rows on each measure, the resilience scores and the well-being outcomes, show the mean values and standard errors of those measures for the groups classified as resilient and non-resilient, and the p-value of the t-test of the comparison of those two distributions.

The results show that the CB method is, as expected, internally consistent by construction. The average values of resilience score between resilient and non-resilient groups are indeed significantly different between each other. This finding is not surprising, because the binary classification in CB is tied to the resilience measure itself. Nevertheless, an interesting result is that the difference in resilience score is maintained also when the probability of well-being refers to a different well-being outcome. In other words, the probability of being above the poverty line is also significantly different between groups that are classified according to their probability of being food secure, and vice versa. The bottom two rows of Table 13 compare the mean values of FCS and RAEC of the resilient and non-resilient groups classified according to the CB method. The results indicate that standard well-being measures are also significantly different between those groups, resilient households having, in average, significantly higher values of FCS and RAEC when classified as resilient by both the conditional probability of being food secure and the conditional probability of being non-poor.

Table 13: t-tests (Resilience Score) between resilient/non-resilient groups, CB

	Resilience classification by: FCS			**Resilience classification by: RAEC**		
	Non-Res.	Res.	t-test (p-value)	Non-Res.	Res.	t-test (p-value)
CB[FCS]	45.10 (0.31)	77.91 (0.25)	0.00	57.51 (0.47)	68.37 (0.49)	0.00
CB[RAEC]	33.12 (0.51)	44.66 (0.56)	0.00	22.43 (0.25)	59.58 (0.36)	0.00
FCS	35.82 (0.39)	47.29 (0.40)	0.00	39.85 (0.40)	44.66 (0.43)	0.00
RAEC	2840.66 (50.17)	3673.55 (55.99)	0.00	2431.27 (33.61)	4209.99 (64.06)	0.00
Observations	1,423	1,627		1,577	1,479	

NOTE: In each row, mean values of resilience score and well-being measures for resilient/non-resilient groups. Standard errors in parenthesis.

We present the results of the t-tests for the RIMA/TANGO binary classification method in Table 14 (recall that the two share the same way of identifying resilient households). Since the binary classification of resilience is also tied here, although in a different way, to the well-being outcomes, we have two resilient and non-resilient groups, one under the classification done for each well-being outcome of interest. For example, the resilient group under the FCS is comprised by the households that did not register a loss in FCS between the second and third survey waves. The resilient group classified by the RAEC is formed by households that did not register a loss in total consumption expenditure between the same time period. Note that, under both RIMA and TANGO, the resilience capacity measures are unitless indexes that are not necessarily directly related by construction to a particular well-being outcome. Our results indicate that under the TANGO/RIMA binary classification method, resilient and non-resilient groups do not have significantly different RCI values. Those results are maintained for the classification done with both the FCS and the

RAEC. This indicates a structural inconsistency between the two distinct measures used under each method, since a higher RCI should be associated with a significantly higher likelihood of being classified as resilient. The fact that these measures are not internally consistent serves as a caution for users.

Lastly, the last two rows in Table 14 present, as in Table 13, the differences in standard well-being measures between resilient and non-resilient groups classified by the RIMA/TANGO binary classification method. In this case, contrary to what we saw with the RCI measures, both groups have significantly different values of FCS and RAEC. This result is not surprising considering that the binary resilience classification under RIMA/TANGO is tied to fluctuations across waves in those same well-being measures. Nevertheless, this result underlines again the dissociation between the resilience capacity measures (RCIs) of the TANGO and RIMA methods with standard well-being outcomes that the results of other analyses presented previously in this thesis have shown (e.g., differences in ranking of households between the RCIs and well-being outcomes).

Table 14: t-tests (RCIs) between resilient/non-resilient groups, TANGO/RIMA

	Resilience classification by: FCS			**Resilience classification by: RAEC**		
	Non-Res.	Res.	t-test (p-value)	Non-Res.	Res.	t-test (p-value)
RIMA[RCI]	23.17 (0.28)	23.37 (0.29)	0.63	23.25 (0.26)	23.32 (0.32)	0.87
TANGO[RCI]	20.65 (0.22)	21.34 (0.24)	0.38	20.96 (0.22)	21.03 (0.25)	0.82
FCS	35.11 (0.33)	48.74 (0.39)	0.00	40.32 (0.35)	44.31 (0.46)	0.00
RAEC	3105.58 (45.91)	3493.24 (54.25)	0.00	2608.06 (32.10)	4303.023 (65.79)	0.00
Observations	1,820	1,879		2,107	1,464	

NOTE: In each row, mean values of RCI and well-being measures for resilient/non-resilient groups. Standard errors in parenthesis.

5.5 Well-being predictive accuracy

5.5.1 Binary resilience classification and normative well-being status

For both RIMA and TANGO, resilience is ”the capacity that ensures that shocks and stressors do not have long-lasting development consequences” (Smith and Frankenberger, 2018; FAO, 2016). For CB, resilience is ”the capacity over time of a person, household or other aggregate unit to avoid poverty in the face of various stressors and in the wake of myriad shocks” (Cissé and Barrett, 2018). Consequently, through time, the well-being of households classified as resilient should be consistent with meeting development goals in the first case, and escaping poverty, however defined, in the second. Both development goals and poverty status are most frequently assessed through normative well-being thresholds. Therefore, a natural analysis is to determine how households that are classified as resilient or not by each one of the methods falls under normative levels of well-being in consecutive time periods. Additionally, this exercise also serves as an example on the correspondence between a targeting based in the binary resilience classification proposed by each one of the methods, and one done by standard well-being thresholds.

We present in Figure 4 a graphic representation of the probability of being above or below a normative well-being threshold in the second and third survey waves conditional on being classified as resilient or not in the second wave. The binary resilience classification criteria used here is the same as the one used for the internal consistency checks – RIMA/TANGO depending on their change in well-being between the first and second waves. CB, probability of being food secure above 65% and of being non-poor above 42% in the second wave. Ideally, to agree with the theoretical definition of resilience, people that are trapped under normative levels of well-being between the two rounds (bottom-left quadrant in each graph) should not be classified as resilient. Conversely, we might expect that the majority of the households that are above normative well-being thresholds in the two periods (upper-right quadrant) are classified as resilient by all the methods. The interpretation of the other two quadrants, bottom-right and upper-left, is not that straightforward with just one round of data. Fluctuations above and below well-being thresholds are commonly found in the development literature (Baulch and Hoddinott, 2000) but their implications for the well-being classification of households is not straightforward. The impossibility of determining with this analysis the nature of those fluctuations, stochastic or structural, complicates the interpretation of these results (Barrett, Garg and McBride, 2016).

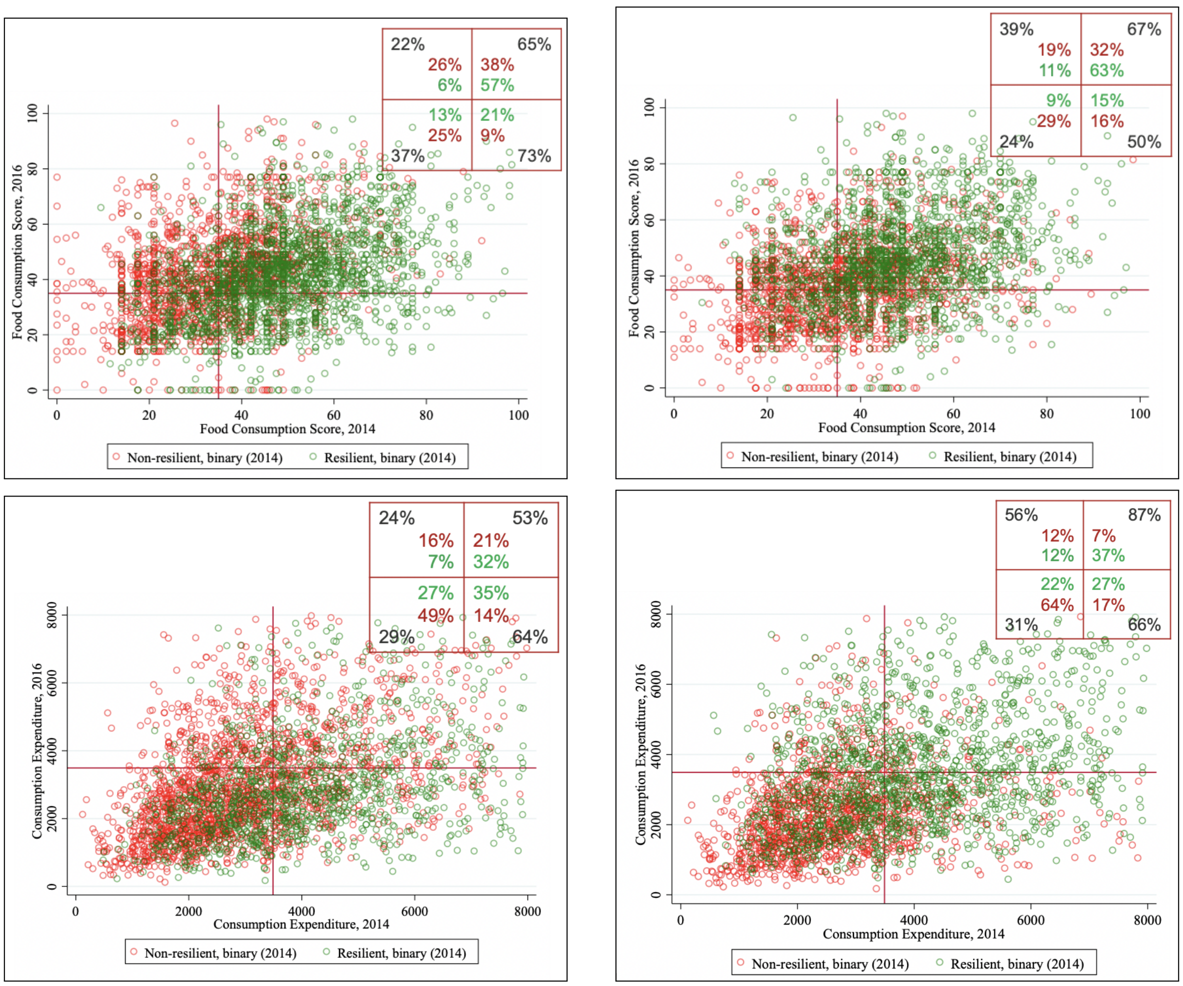

Figure 4: Binary resilience and normative well-being. RIMA/TANGO [left]; CB [right]

In Figure 4, the black numbers inside each box refer to the percentage of the total observations in that quadrant classified as resilient. The green numbers refer to the percentage from the total resilient households in the sample that are inside of that quadrant, and the numbers in red the percentage from the total non-resilient households in the sample that fall in that quadrant.

The results on how resilient and non-resilient households fall under normative well-being levels differ between resilience classification methods and well-being outcomes. When FCS is the outcome of interest, most of the resilient households are correctly classified in the upper-right quadrant in both RIMA/TANGO and CB binary classification methods, and most of the food-insecure households in both waves (bottom-left quadrant) are classified as non-resilient. Under CB, 67% of the households that are food secure in both waves -2014 and 2016- were classified as resilient in 2014, and 76% of the households in the bottom-left quadrant, the ones that are food insecure in both waves, are classified as non-resilient. Under the RIMA/TANGO method, those same percentages go down to 65% and 63%, respectively. We also look at how resilient and non-resilient groups are classified across quadrants. The results indicate that just 29% of all the non-resilient households under the CB method and 25% under RIMA/TANGO are food insecure in both periods. A great percentage of the households classified as non-resilient – 32% in CB and 38% in RIMA/TANGO – are food secure in both waves. This implies that, if these classification methods were used for targeting, we might be focusing on households that are not that worse off -at least according to the FCS-. Recall that, in our sample, almost 50% of the units in wave two are considered non-resilient, so that group is comprised by a significant number of households.

Looking at the results for the consumption expenditures, the results are somewhat the opposite as what we found for the FCS. Now, the greatest percentage of the non-resilient households falls correctly in the bottom-left quadrant, in both CB and RIMA/TANGO methods – 64% of all non-resilient households are in the bottom-left quadrant in CB and 49% in RIMA/TANGO –. However, resilient households are now more dispersely classified across quadrants, both in CB and RIMA/TANGO. In CB, just 37% of the resilient households are non-poor in both waves, and 32%

in RIMA/TANGO. Almost 25% of all the households classified as resilient are below the poverty line in both waves for both methods. This misclassification can have worst consequences that the one that we saw in the FCS, when households classified as non-resilient were food secure in both waves, because if this classification method is used for targeting purposes, those households classified as resilient might not be targeted, even though they fall under normative well-being levels in consecutive waves. In our sample, 42% of the households are considered resilient according to the RAEC, so that implies that, with a sample size of 3,776 households, circa 350 households in CB and 430 in RIMA/TANGO are classified as poor and resilient.

Finally, when interpreting those results, it should be taken in account that the probability cut-off that classifies households as resilient or not in the CB method is arbitrarily chosen. This implies that different, more restrictive probability cut-offs could be selected, specially in the case of consumption expenditures, and see how the percentage of resilient households under each quadrant changes.

Note that in Figure 4 the percentages of the relative classification of resilient and non-resilient households in each quadrant, represented by the green and red numbers in the boxes, should add up to a hundred. However, when FCS is the well-being outcome of interest, this is not the case, due to some households that fall in the limiting FCS score of thirty-five in one wave and not in the following one, or vice-versa. Those households cannot be classified in an specific quadrant and are therefore left out of the classification (around 4% of households). For example, the first quadrant (bottom-left) contains households that have values of FCS that are equal or lower than thirty-five in both waves. None of the other quadrants contain households that fall in the limiting score. If a household has a score equal to that limiting value in one wave and not in the other, it will be left out of the classification. This dies not happen in the case of the consumption expenditures because we do not have households that have values on consumption expenditures equal to the poverty line.

5.5.2 Out-of-sample well-being predictive accuracy

Resilience capacity does not necessarily have to be completely associated with well-being outcomes. If that were the case, we will be just mimicking already existent measures, without incorporating anything new from a development perspective. However, resilience cannot be completely dissociated from well-being either, when the ultimate goal is to increase and/or stabilize the well-being of the targeted communities in the short and long run. Therefore, an interesting comparative exercise is to determine to what extent each one of the methods better predicts current and future well-being states. For each method, we estimate the well-being predictive accuracy of the models that relate well-being outcomes and resilience capacity measures (equation 11, page 17), and of the resilience capacity measures themselves (equation 12, page 18). We do both exercises in cross-sectional and panel forms, so as to determine predictive accuracy of both current and future well-being states.

We present first the results of the well-being predictive accuracy of the resilience capacity measures of each one of the methods (Tables 14 and 15). We do the exercise in both cross-sectional and panel forms; estimating the measures in a 75% random sample of the third survey wave and predicting in the 25% left for the cross-sectional case; and estimating the measures with all-but-last survey waves in the panel case, predicting in the last survey wave. In both approaches, we run an OLS regression of real well-being over the predicted resilience capacity measures. Since well-being outcomes and resilience capacity measures are not in the same units, we report model fit through the RMSE of each regression. The RMSE is a measure of absolute model fit, that represents the square root of the variance of the residuals. The lower the RMSE, the better the model fit. Its units are the same as the ones of the model's dependent variable, in our case values of FCS and consumption expenditures.

Agreeing with what we saw in previous comparative analyses, these results indicate that CB's resilience score is a closer representation of well-being status, predicting well-being more accurately (lower RMSE) than the other two measures in both the cross-sectional and panel cases. RIMA and TANGO have slightly worst results that are very similar to one another. This result is not surprising, taking in account that CB's resilience score is a measure directly associated to well-being by construction. Note that in the CB method, we use just the resilience score generated for that specific well-being outcome (e.g., to predict current and future FCS, we use the resilience score that represents the conditional probability of being above the FCS well-being threshold). We additionally include in the panel case the results of the well-being prediction of lagged well-being outcomes. Our results indicate that the prediction of well-being by lagged well-being measures is very similar to that predicted by CB's resilience scores. These results indicate that there is not gain from using the resilience capacity measures of both the RIMA and TANGO methods to predict well-being across waves.

To show those differences in predictive accuracy more clearly, we present in Figure 5 a vertical rule were we compare the predictive accuracy of lagged well-being with the three resilience capacity measures. The graph on the left shows the FCS predictive accuracy of lagged FCS and the three resilience capacity measures, and the graph on the right the predictive accuracy of the consumption expenditures by the lagged well-being measure and the three resilience capacity measures. In both graphs, the solid horizontal line represents the well-being predictive accuracy of lagged well-being. The results show how both RIMA and TANGO measures of resilience capacity do not represent any gain in predictive accuracy compared to simple lagged well-being, while the use of CB's resilience score entails an increase in predictive accuracy when compared to lagged well-being.

Finally, in the cross-sectional case, we do several draws and generate a distribution of RMSEs, which allows us to perform pair comparisons of means to see if the values of RMSE are indeed different between models (Table 17). The results indicate that the RMSEs are significantly different between models in all cases, except between TANGO and RIMA models when FCS is the outcome of interest.

Table 15: RMSEs from cross-sectional well-being prediction by resilience capacity measures

	TANGO	**CB**	**RIMA**
FCS	16.42 (0.39)	15.38 (0.41)	16.41 (0.38)
RAEC	2051.95 (111.36)	1824.62 (104.96)	1961.16 (94.54)
NOTE: Standard errors in parenthesis. Estimation and prediction in wave 3			

Table 16: RMSEs from panel well-being prediction by resilience capacity measures

	TANGO	**CB**	**RIMA**	**FCS_{t-1}**	**$RAEC_{t-1}$**
FCS	16.33	15.11	16.64	15.92	16.71
RAEC	2070.28	1831.82	2103.11	2103.02	1915.91
NOTE: Estimation in all-but-final survey waves, prediction in last wave					

Table 17: p-values of pair comparison of cross-sectional RMSEs between models

	TANGO-CB	**TANGO-RIMA**	**CB-RIMA**
FCS	0.00	0.38	0.00
RAEC	0.00	0.00	0.00
NOTE: Distribution of the RMSEs from 10 random draws of cross-sectional data.			

We also test the predictive well-being accuracy of the specifications that relate resilience capacity and well-being outcomes in each model. We do it again both cross-sectionally and in panel forms, following the same procedure as for the resilience measures but now estimating those specifications -in RIMA the SEM-MIMIC model (eqs. 1 and 2, pg 7), in TANGO the cross-sectional regression model (eq. 4, pg 9) and in CB the main regression model (eq. 8, pg 12)-. Then we run a regression of real and that predicted well-being (eq. 11, pg 17). Tables 18 and 19 show the results of the intercept and slope of those regressions, for the cross-sectional and panel cases, respectively. Again, CB's model that includes the well-being path dynamics is the one that yields the closest results, specially in the cross-sectional case, with intercept estimates that are closer to zero and slope coefficient estimates closer to one. In this exercise, RIMA's SEM-MIMIC model is

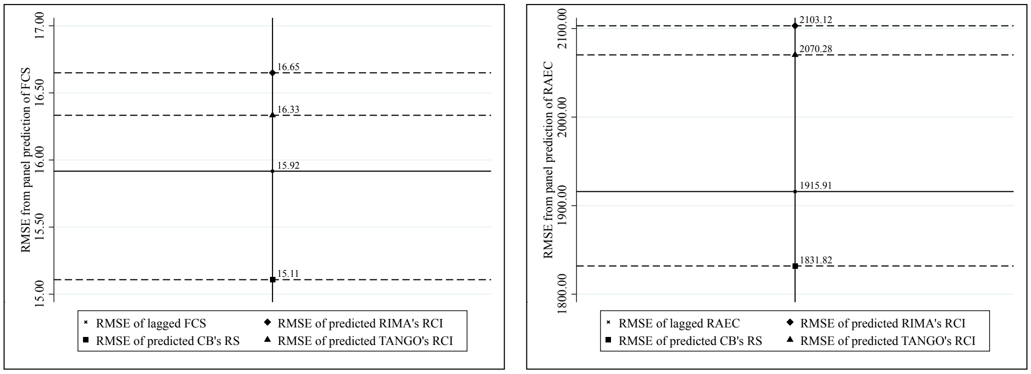

Figure 5: Panel well-being predictive accuracy of lagged well-being vs. resilience cap. measures

the one that performs the worst. Although the authors specify that the MIMIC model is in no case used to determine a causal relationship between resilience capacity and well-being (D'Errico et al., 2018), these results indicate that the formalization of that relationship under that specification is not as accurate as other available statistical approaches.

Finally, we test with Wald tests the composite linear hypothesis of the slope being equal to one and the intercept to zero. Table 20 presents the F-statistics of those tests, both for the cross-sectional and panel cases for each one of the three methods. The results indicate that the joint hypothesis cannot be rejected for the cross-sectional well-being prediction by TANGO and CB models. Nevertheless, for those same two models, the hypothesis is rejected when the well-being prediction is through time (panel prediction). In the RIMA case, the joint hypothesis is always rejected, both in the cross-sectional and panel approaches.

Table 18: Intercept and slope estimates from cross-sectional well-being prediction

	TANGO		**CB**		**RIMA**	
	Intercept	Slope	Intercept	Slope	Intercept	Slope
FCS	-4.11	1.05	0.65	0.96	-28.84	1.52
	(1.08)	(0.02)	(1.29)	(0.02)	(1.65)	(0.03)
RAEC	-273.33	1.11	-189.87	1.07	-1671.25	1.54
	(119.56)	(0.04)	(75.14)	(0.02)	(59.06)	(0.02)

NOTE: Standard errors in parenthesis. Estimation and prediction in wave 3

Table 19: Intercept and slope estimates from panel well-being prediction

	TANGO		**CB**		**RIMA**	
	Intercept	Slope	Intercept	Slope	Intercept	Slope
FCS	6.86	0.82	11.25	0.90	-27.37	1.57
RAEC	-531.79	0.96	-96.13	0.72	-2358.38	1.57

NOTE: Estimation in all-but-final waves, prediction in last wave.

Table 20: F-statistics of joint hypothesis (slope=1; intercept=0), Wald tests

	TANGO		**CB**		**RIMA**	
	cross-sectional	panel	cross-sectional	panel	cross-sectional	panel
FCS	0.23	0.00	0.62	0.00	0.00	0.00
RAEC	0.48	0.00	0.61	0.00	0.00	0.00

5.6 Resilience capacity as a buffer between shocks and well-being

So far, we have seen that each one of the methods differs on their approach on how to conceptualize and measure resilience capacity empirically. One question that arises naturally from these differences in variable selection and statistical approach is which one of the three versions captures better the concept of resilience, as the capacity that helps households to buffer the negative effects of a shock on well-being. We address this by estimating equation (13) (pg.19), for each one of the methods.

Table 21 presents the results of those regressions. All the regressions include controls, but they are not presented here for simplicity (for a complete table with all the variables consult Appendix G). The results indicate that in both RIMA and CB, resilience capacity measures are positively and significantly correlated with the well-being outcomes. The coefficient is also positive in the case of TANGO, but just weakly significant when consumption expenditures is the outcome of interest. The magnitudes of the significant coefficients, those of RIMA and CB, are very similar, indicating similar impacts of the resilience capacity measures on well-being.

The results of the measures of drought and interaction terms are puzzling. The measures of drought appear to have a positive effect in the outcome variables, which might reflect that the measure of drought used in the regressions does not capture adequately the negative effects of drought on well-being.

To interpret these results it is important to take into account some facts about the nature of the data and the resilience measures themselves. First, the total number of observations in our sample that are affected by a drought – that have a mean normalized NDVI below -1.5 SD – is very limited. From the more than 3,500 households in each wave, just around 250 are affected by a drought. This restricts the total variability in resilience capacity for the households affected by drought, which might make it more difficult to capture the buffering effect in that coefficient. Additionally, the generation of the truncated variable might induce to an indirect non-random selection of the sample by other attributes rather than exposure to shock. Additionally, the estimation of the CB resilience score requires more waves of data due to the inclusion of lagged well-being in the main model. As a result, we just have CBs resilience capacity measures for two waves, and since these equations include resilience capacity in lagged form, CBs regression has just half the observations of the other two methods. This restricts even more the variability in the sample, which makes it more difficult to capture any effects in a regression context.

Additionally, as it was mentioned in previous sections of this document, the resilience measures of the three methods do not measure exactly the same concepts, which complicates the interpretation of the results of this comparison exercise. While in both TANGO and RIMA the resilience measure is a unitless index, in CB is the probability of being above a certain normative well-being threshold. All those differences in sample size and meaning of the resilience measures make it difficult to determine what is the adequate comparison benchmark to evaluate how these measures should behave in a regression context, specially between TANGO/RIMA and the CB measures.

Lastly, motivated by these counterintuitive results, the estimation of the regression model was repeated using an alternative measure of drought. Despite the risk of incurring in measurement error, I included in the estimation the self-reported measure of drought provided by the LSMS-ISA Ethiopian data. The results of that alternative analysis are not presented in this document, since they remain qualitatively unchanged.

Table 21: Buffering capacity of the resilience capacity measures

	FCS	RAEC	FCS	RAEC	FCS	RAEC	FCS	RAEC
RCI [RIMA], (t-1)	0.14***	30.52***						
	(0.04)	(4.17)						
RCI [RIMA], (t-1)*Drought	-0.09	-1.39						
	(0.07)	(4.72)						
RCI [TANGO], (t-1)			0.12*	12.98				
			(0.08)	(8.67)				
RCI [TANGO], (t-1)*Drought			-0.42***	-2.86				
			(0.13)	(11.92)				
RS [CB, FCS], (t-1)					0.25***	22.11***		
					(0.04)	(4.64)		
RS [CB, FCS], (t-1)*Drought					0.15***	-1.30		
					(0.04)	(3.95)		
RS [CB, RAEC], (t-1)							0.14***	26.28***
							(0.04)	(3.28)
RS [CB, RAEC], (t-1)*Drought							-0.07	-1.98
							(0.07)	(4.14)
Drought	12.17*	747.10*	21.23***	813.40	-9.97**	617.00	8.05	632.10*
	(6.15)	(420.00)	(6.16)	(521.50)	(6.87)	(625.60)	(8.02)	(360.40)
Drought^2	-4.19*	-353.00**	-5.84***	-378.90**	0.51	-274.50	-2.42	–263.60*
	(2.28)	(175.40)	(2.03)	(187.70)	(2.46)	(199.40)	(2.62)	(135.40)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	7,274	7,046	6,768	6,556	3,187	3,082	3,074	2,996
R-squared	0.079	0.208	0.084	0.186	0.156	0.256	0.075	0.267

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

CHAPTER 6

SUMMARY AND CONLCUSIONS

Our comparative exercise has shown that there are substantial differences between these three widespread resilience measurement methods, and that there is still space for improvement in all of them. When estimating resilience capacity, households in the same sample are ranked differently between methods, specially between CB's resilience score and RIMA and TANGO resilience capacity indexes. The methods therefore differ in the core of any resilience analysis, the empirical identification of resilience capacity. The classification of the ”worst off” according to resilience capacity is also very different between them. If that measure were used for targeting, the selected group of households will substantially differ between methods.

Resilience capacity measures and familiar well-being outcomes such as real expenditures per adult equivalent or the food consumption score are strongly correlated under the CB method than with the other two resilience measurement methods. CB's resilience score is a closer representation of well-being by construction since it is generated as a conditional probability of achieving a given level of a well-being measure. In the case of RIMA and TANGO, resilience measures are fairly dissociated from well-being, making it difficult to determine how a potential increase in resilience capacity can affect well-being status or which population groups are being represented under their resilience capacity measures in terms of well-being.

A key aspect of the resilience empirical analysis, the identification of resilient and non-resilient households, indicates a lack of consistency between resilience capacity measures and the binary classification under RIMA/TANGO method. Note that both agencies, the FAO and TANGO, have not positioned themselves on how the targeting will be done. However, this result indicates a complete dissociation between both approaches for identifying resilience empirically, which should serve as a warning for practitioners when using those measures. In the CB method, since the definition of resilient units is tied to the resilience capacity measure, resilient and non-resilient groups differ significantly in their average resilience score values.

The out-of-sample well-being predictive accuracy of the binary resilience classification shows promising results. Most of non-resilient households fall indeed under normative levels of well-being through time. Even though resilience capacity should not be an identical representation of well-being, analyzing fluctuations on normative well-being thresholds according to resilience status is directly derived from the definition of resilience itself and should be taken into account in the empirical resilience analysis.

The out-of-sample well-being predictive accuracy of the resilience capacity measures differs between methods. Consistent with other results, CB's resilience score outperforms the other two measures in predicting well-being outcomes out-of-sample, both in cross-sectional and in panel prediction. In the case of panel prediction, CB's is at least as accurate as lagged well-being outcomes. Conversely, our results indicate that there is no gain from predicting well-being through time with the resilience capacity measures of the RIMA and TANGO methods, compared to just using lagged well-being. This result underlines the disconnection between those RCIs and standard well-being outcomes. Resilience capacity, by definition, should be somehow related to well-being, specially trough time. If those resilience capacity indexes specified as they are now cannot capture fluctuations in well-being for the analyzed population, it might be the case that the theoretical premises that lead to their estimation should be revised. Perhaps the selection of variables, informed by theory, is not adequate in this context, or the way on which the variables are combined through data reduction techniques is not appropriate for this dataset. In this still ongoing process of developing robust and empirical applicable resilience measures, these results should be taken as a warning for practitioners.

Finally, despite differences in predictive accuracy and ranking, none of the measures passes the test of the resilience capacity measure as a buffer of the negative effects of drought on well-being. The results of this exercise are nevertheless difficult to interpret, since the measures do not represent exactly the same concepts, which complicate the identification of a common comparative benchmark in a regression context, specially between TANGO/RIMA and the CB measures. Furthermore, we do not know the true value, even sign, of the relationship between the well-being measures and the interaction term, so interpreting this test poses challenges given that we do not know which measure gets us closer to an unknown true value.

This first estimation of the three resilience methods with a common dataset has yielded interesting comparative results, mostly in terms of caution when applying those methods empirically to guide or evaluate development interventions. However, there are some caveats or limitations in our analysis that should be taken in account.

The first one is the limited identification of shock exposure that we have in our sample. Very few of the observations were actually affected by a shock according to our measure of drought, which limits the possibilities of focusing on just that population sub-group. We can imply that most households in this context might be virtually affected by several types of shocks, and therefore changes in well-being can be associated with those. However, in the sake of robustness, more information on how resilience capacity and well-being are related in the households for which being affected by a shock can be confirmed will enrich our analysis.

Second and last, we have not explored the different weight that every component of resilience capacity has on determining the final measure. This is an important caveat, currently not empirically solved, that related to the difficulties that development agencies have in determining how best to implement resilience capacity building interventions. Future research efforts can be oriented to fill this gap in impact evaluation.

Summarizing, even with the limitations that we just pointed out, there are some key takeaways derived from our analysis empirical analysis of resilience measures. First, resilience capacity should be somehow related to well-being, and determining which population groups in terms of well-being are being represented under different values of resilience capacity is crucial from a development perspective. Second, the procedure to target households in terms or resilience is still unclear, which complicates the work of development agencies. Third, different resilience measures yield different values of resilience capacity for the same households, which implies that, at the moment, the work of different development organizations that use different methods is not directly comparable. Finally, determining which characteristics are important to build resilience capacity is a crucial aspect of the resilience analysis that has been not yet explored, and effort should be put towards filling this gap.

BIBLIOGRAPHY

Abay, K.A., Abate, G.T., Barrett, C.B., Bernard, T. (2019). Correlated non-classical measurement errors, Second best policy inference, and the inverse size-productivity relationship in agriculture. *Journal of Development Economics, 139*(2019): 171-184

Alloush, Mo (2019). Resilience and Psychological Well-being. Dissertation Chapter, Univesrity of California at Davis.

Alinovi, L., Mane, E. and Romano, D. (2008). Towards the measurement of household resilience to food insecurity: An application to Palestinian households. In Sibrian (Ed.),*Deriving food security information from national household budget surveys. Experiences, achievement, challenges (pp. 137-152)*. Rome, FAO.

Alinovi, L., d'Errico, M., Mane,E. and Romano D.(2010). Livelihood strategies and household resilience to food insecurity: an empirical analysis to Kenya. Background paper to the European Development Report 2010. Fiesole: European University Institute.

Barrett, C. B. and Constas, M. A. (2014). Toward a Theory of Resilience for International Development Applications. *Proceedings of the National Academy of Sciences 111*(40): 14625-14630.

Barrett, C. B., Garg, T. and McBride, L. (2016). Well-Being Dynamics and Poverty Traps. *Annual Review of Resource Economics 8*: 303-327.

Baulch, B. and Hoddinott, J.F. (2000). Economic mobility and poverty dynamics in developing countries.*Journal of Development Studies,36 (6)*(2000): 1-24

Béné, C., Wood, R.G., Newsham, A. and Davies, M. (2012). Resilience: New Utopia or New Tyranny? Reflection about the Potentials and Limits of the Concept of Resilience in Relation to Vulnerability Reduction Programmes. Institute of Development Studies (IDS) working paper, Vol. 2012 Num.405. Centre for Social Protection (CSP) working paper, Num. 006

Cameron, A.C. and Miller, D.L. (2015). A Practitioners Guide to Cluster-Robust Inference. *J. Human Resources* 2015 vol. 50 no. 2 317-372

Cissé, J. D. and Barrett, C.B. (2018). Estimating development resilience: A conditional moments-based approach. *Journal of Development Economics, 135*(2018): 272-284.

Cissé, J. D. and Ikegami, M. (2016). Does insurance improve resilience? Measuring the impact of index-based livestock insurance on development resilience in Northern Kenya.

Constas, M.A., Frankenberger, T. and Hoddinott, J.F. (2014). Resilience measurement principles: Toward an agenda for measurement design. Resilience Measurement Technical Working Group Technical Series 1. Rome: Food Security Information Network.

Constas, M.A., Frankenberger, T. and Hoddinott, J.F., Mock, N., Romano, D., Béné, C. and Maxwell, D. (2014). *A Common Analytical Model for Resilience Measurement.*Resilience measurement technical Working Group. Technical series no. 2. Rome: Food Security Information Network.

Dercon, S. and Krishnan, P. (1998). *Changes in Poverty in Rural Ethiopia 1989-1995: Measurement, Robustness Tests and Decomposition*. Centre for the Study of African Economies. Institute of Economics and Statistics. Oxford, UK.

d'Errico, M., Garbero, A., Constas, M.A. (2016). *Quantitative Analyses for Resilience Measurement. Guidance for constructing variables and exploring relationships among variables*. Resilience measurement technical Working Group. Technical series no. 7. Rome: Food Security Information Network.

d'Errico, M., Romano, D. and Pietrelli, R. (2018). Household resilience to food insecurity: evidence from Tanzania and Uganda. *Food Security 10(4)*: 1033-1054.

FAO (2016). Reseilience Index for Measurement and Analysis II: Analysing resilience for better targeting and action. Food and Agriculture Organization of the United Nations, 2016, Rome.

Frankenberger, T. and Smith, L. (2016). Ethiopia Pastoralist Areas Resilience Improvement and Market Expansion (PRIME) Project Impact Evaluation: Report of the Interim Monitoring Survey 2014-2015, DRAFT. February 2016. Washington, DC: USAID Feed the Future.

Garbero, A. and Chichaibelu, B.B. (2016). "Estimating resilience outcomes in an impact assessment framework with high-frequency data". 2018 Annual Meeting, August 5-7, Washington, D.C. 274460, Agricultural and Applied Economics Association.

Henly-Shepard, S. and Sagara, B. (2018). Resilience Measurement Practical Guidance Note Series: An Overview. Produced by Mercy Corps as part of the Resilience Evaluation, Analysis and Learning (REAL) Associate Award.

Hu, L. and Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1-55.

Knippenberg, E., Jensen, N. and Constas, M. (2019). Quantifying household resilience with high-frequency data: Temporal dynamics and methodological options. *World Development, 121*(2019): 1-15

Leroy, J.L., Ruel, M., Frongillo, E.A., Harris, J. and Ballard, T.J. (2015). Measuring the Food Access Dimension of Food Security: A Critical Review and Mapping of Indicators. *Food Nutrition Bulletin 36*(2)(2015):167-95.

Naude, W., Santos-Paulino, A. and McGillivray, M.(2009). Measuring Vulnerability: An Overview and Introduction. *Oxford Development Studies, 37*, 183191.

Phadera, Lokendra, Hope Michelson, Alex E. Winter-Nelson, and Peter D. Goldsmith (2019). Do asset transfer build household resilience? *Journal of Development Economics, 138*: 205-227.

Ravallion, M. (1992). Poverty Comparisons: A Guide to Concepts and Methods. Living Standards Measurement Study Working Paper N.88. World Bank, Washington D.C.

Smith, L. and Frankenberger, T. (2018). Does resilience capacity reduce the negative impact of shocks on household food security? Evidence from the 2014 floods in Bangladesh. *World Development 102*(2018):

358-376.

TradingEconomics (2019). Ethiopia-Economic Indicators. Available online at: https://tradingeconomics .com/ethiopia/indicators. (Last visit: June, 2019).

United Nations (2018). *Ethiopia's progress towards eradicating poverty*. Paper to be presented to the Inter-Agency Group Meeting on the "Implementation of the Third United Nations Decade for the Eradication of Poverty (2018-2027)". United Nations Development Program (UNDP).

Upton, J. B., Ciss, J.D., and Barrett, C.B. (2016). Food security as resilience: reconciling definition and measurement. *Agricultural Economics, 47*(2016): 135-147.

Weismann, D., Bassett, L., Benson, T., Hoddinott, J.F. (2009). Validation of the World Food Programmes Food Consumption Score and Alternative Indicators of Household Food Security. IFPRI Discussion Paper 00870, June 2009. International Food Policy Research Institute, Washington D.C.

World Bank (2019_a). Living Standards Measurement Surveys: Microdata Catalogue. Available on line at: http://microdata.worldbank.org/index.php/catalog/lsms (Last visit: June, 2019).

World Bank (2019_b). Living Standards Measurement Surveys: Programs:LSMS-ISA:Ehtiopia. Available on line at: http://surveys.worldbank.org/lsms/programs/integrated-surveys-agriculture-ISA/ethiopia (Last visit: June, 2019).

APPENDIX A

DEFINITION OF RIMA PILLARS AND TANGO RESILIENCE CAPACITIES

In this Appendix, we present the theoretical intuition behind the variable composition of the resilience pillars under the RIMA method and the resilience capacities for TANGO, derived from the original papers of each method (FAO, 2016; Smith and Frankenberger, 2018). With this, we want to help to the understanding of the original motivation behind the inclusion of an specific set of resilience capacity variables proposed by the authors of each one of those two methods.

For RIMA, the theoretical definition of what the different pillars want to capture is the following,

1. **Access to Basic Services (ABS):** measured by the presence of different services in the community, and potentially also in terms of quality by the total cost of those services or subjective measures. The role of this pillar in building resilience is justified by its effect on generating income from assets, determining risk exposure and, in an indirect way, somehow reflecting the relationship between governmental institutions and the general population.
2. **Assets (AS):** productive and non-productive assets are included based on their previously reported importance in determining consumption smoothing capacity and coping strategies in the face of shocks.
3. **Social Safety Nets (SSN):** identified by informal and formal transfers, as an indirect measure of social cohesion. Important in terms of alleviating the consequences of a shock, like assets, increasing consumption smoothing capacity.
4. **Sensitivity (S):** refers to the measure of exposure to different risks, both in terms of frequency and intensity. The inclusion in the pillars is in terms of further analyzing how shocks affect resilience capacity.
5. **Adaptive Capacity (AC):** it reflects the capacity of reorganize and adapt to a changing environment. RIMA-II proposes its identification with different measures of human capital, such as total education attainment or disability; income source and agricultural diversification and household composition related variables, such as the dependency ratio.

In the TANGO case, the theoretical definition of the pillars are presented below.

1. **Absorptive Capacity:** ”ability to minimize exposure to shocks and stressors and to recover quickly when exposed”. Composed by: bonding social capital, asset ownership, cash savings, access to informal safety nets and disaster preparedness and mitigation.
2. **Adaptive Capacity:** ”making proactive and informed choices about alternative livelihood strategies”. Composed by: bridging social capital, linking social capital, aspirations and confidence to adapt, livelihood diversity, asset ownership, human capital, exposure to information and access to financial services.
3. **Transformative Capacity:** ”enabling conditions that foster more lasting resilience”. Composed by: bridging social capital, linking social capital, access to markets, access to services, women’s empowerment, quality of governance and access to formal safety nets.

APPENDIX B

FACTOR ANALYSIS IN THE TANGO METHOD

TANGO uses an unorthodox variation of the regular principal factor analysis method to estimate the multiple indexes needed to generate its resilience capacity measure. We follow that procedure to estimate the indexes used to replicate the TANGO method. The steps of that procedure are the following,

1. Factor loadings and factor scores for each index are estimated using principal factor analysis with data from the first survey wave. The results are manually screened by the researcher, and any variables that have a KMO lower than 0.50 or that have the ”wrong sign” according to the ”consistency with the meaning” of the concept that the factor wants to capture are dropped out of the estimation of the final index (Frankenberger and Smith, 2016 (pg.12); Smith and Frankenberger, 2018 (pg.362); and internal document prepared by TANGO international). The re-estimated factor loadings of the remaining selected variables are stored. Note that best practice in factor analysis would typically use the full set of data analyzed, not just the first survey wave, and would not drop variables with factor loadings subjectively assessed to have the wrong sign. Screening using KMO statistics is common, albeit often frowned upon by experts.

2. In the subsequent waves of data, the selected variables are standardized, according to the mean and standard deviation of those same variables in the first wave. In other words, the variables that will form part of the index are expressed in other waves as standard deviations from the first wave's mean. Mathematically,

$$Standardized\ Var_{other\ waves} = Var_{other\ waves} - Var\ Mean_{wave\ 1} / Var\ SD_{wave\ 1}$$

3. The stored factor loadings of the finally selected variables, that were re-estimated after that initial screening process, are multiplied by the standardized variables in waves other than the first. The factor analysis is not recomputed for the remaining dataset. Each component is added up to generate final factor index according to the following expression,

$$Factor\ Index = \sum_{i=1}^{n} FS_{wave\ 1} * St\ Var_{wave\ j}$$

where each factor index is a vector of n first wave factor scores (FS) multiplied by the correspondent standardized variable ($St\ Var$) in wave j.

4. The final factor indexes are re-scaled from 0 to 100, using a mini-max procedure according to the following expression, where Min is the minimum value of the estimated index and Max its maximum value,

$$FactorIndex_{0-100} = (FactorIndex - Min)/(Max - Min) * 100$$

APPENDIX C

THE LSMS-ISA ETHIOPIAN DATASET

ETHIOPIA- Living Standards Measurement Study- Ethiopia Socioeconomic Survey (ESS)

Wave 1: 2011/2012

The first survey wave covers all regional states except the capital, Addis Ababa, being representative of rural areas and small towns of Ethiopia. It was originally implemented in 290 rural and 43 small town enumeration areas (EAs). There are two randomization levels: EAs and households (sample size: 12 households for each EA; Total: 3,969 hh). The information is collected in a total of five questionnaires: household (all households in the sample); community (socio-economic indicators of the EAs); post-planting agricultural questionnaire, post-harvest agricultural questionnaire (last two conducted in the same fields and crops) and livestock questionnaire (last three administered to all households engaged in agricultural activities (10 of the total 12 in each EA). Figure C1 shows the timeline of the data collection of each questionnaire in wave 1. Shaded blocks represent month of data collection.

ERSS Questionnaire	**2011**				**2012**		
	Sep	Oct	Nov	Dec	Jan	Feb	Mar
Post-planting agriculture questionnaire							
Livestock questionnaire							
Household, community, and post-harvest agriculture questionnaire							

Figure C1: Data collection time-line, wave 1

Wave 2: 2013/2014

This survey wave was expanded to cover all regional states including the capital, Addis Ababa. It is therefore representative of rural, small towns and urban areas of Ethiopia. The total number of households interviewed increased from 3,776 (333 EAs) to 5,262 (433 EAs) (Note that in our analysis we just use the original sample wihout urban areas). From the 3,969 hh interviewed in the first wave, 3,776 were re-interviewed in this second wave. That equates to an attrition level of 5%. The two levels of randomization are maintained: EAs (rural areas, small town and mid and large towns) and households (sample size: 12

households for each EA; 15 households from each mid and large town EA. The information was collected in the same five questionnaires as in wave 1 (Including some new features in some categories). Figure C2 shows the data collection timing for each questionnaire in the second wave. Again, shaded blocks indicate the months of data collection.

ESS Questionnaire	**2013**				**2014**		
	Sep	Oct	Nov	Dec	Feb	Mar	Apr
Post-planting agriculture questionnaire							
Livestock questionnaire							
Household, community, and post-harvest agriculture questionnaire							
All modules in major urban EAs							

Note: Agriculture interviews for urban areas didn't include crop-cut and land measurement by the enumerator. Agriculture information collected from those EAs in major urban areas was based on self-response.

Figure C2: Data collection time-line, wave 2

Wave 3: 2015/2016

The same levels of randomization and questionnaire structure is maintained in this last wave. From the original 3,969 hh of wave one, 3,699 were re-interviewed in wave three, which corresponds to an attrition level of 7%. Figure C3 shows the months of data collection for each one of the five questionnaires, which coincide with the ones of wave two.

ESS Questionnaire	**2015**				**2016**		
	Sep	Oct	Nov	Dec	Feb	Mar	Apr
Post-planting agriculture and Livestock questionnaires							
Crop Cut							
Household, Community, and Post-harvest agriculture questionnaire							
All modules in large town EAs							

Figure C3: Data collection time-line, wave 3

APPENDIX D

GENERATION OF RESILIENCE VARIABLES. ADAPTATION OF THE ORIGINAL GUIDELINES TO ETHIOPIAN LSMS-ISA DATA.

As mentioned in Chapter 3, one of the key aspects of this paper is to allow for the comparability between the results of the three resilience models, while maintaining fidelity as close as possible to their original guidelines. In terms of variable selection, we adapted the original guidelines of each paper to the information available in the Ethiopian LSMS-ISA datasets. In this appendix, we present how the variables presented in table 4 (pp.28 and 29) are combined to generate the indexes and sub-indexes used to construct the measures in each method. The results are presented in comparative form, to ease the identification of the similarities and differences between the original papers and our own adaptation.

Table D1 shows the comparison of the variable selection and statistical generation procedure for the TANGO method in both the original Smith and Frankenberger (2018) paper and our own work.In their original paper, TANGO's method was estimated with a purpose-built dataset, collected for a certain development program. This is not directly comparable to more standard, LSMS-style datasets. To overcome this difficulty, we match as close as possible the general concepts that TANGO refers to with their variable selection, but adapting it to the LSMS-ISA data. We were unable to match exactly just two of the fourteen original TANGO variables we were not able to do match exactly. All the variables used for the replication of the TANGO indexes referred to in Table D1 can be also found in table 4 of this same document, with the exception of the ones that refer to the presence of certain basic service in the community (EA level). Those correspond to dummy variables, created from the distance variable that is present in the LSMS data. They take a value of 1 if the distance to the service is zero (the service is inside the EA area). This transformation was done to match better TANGO's guidelines.

Much simpler was the adaptation of RIMA to our data. Table D2 shows the mapping of the original and own variable generation of the four pillars. The variables in each category are combined through factor analysis to generate an index for each one of them.

Table D1: Original and own variable selection for estimation of TANGO method

		TANGO (Smith and Frankenberger, 2018)	Adaptation to LSMS-ISA data
Absorptive Capacity	Bonding Social Capital	Additive Index (0-4)	Factor Analysis Index
		(Yes/No) Whether would be able to borrow money or food from friends/relatives under urgent need (Yes/No) Whether would be able to borrow money or food from others in village under urgent need	Prob. of borrowing from close circle % of hh in the community (EA) that received informal transfers Value of informal transfers (hh level)
	Asset Ownership	Factor Analysis Index	Factor Analysis Index
		Domestic assets Productive assets Livestock Land	Durable asset index Productive asset index Livestock Ownership (TLU) Access to agricultural land
	Access to informal safety nets	Factor Analysis Index	Factor Analysis Index
		Existence of women's group (by program) Existence of savings group (village level) % of households at village level that could borrow money from others in village (non-relatives) if needed	% of hh in the community (EA) that received any type of assistance Presence of cooperatives in the community (EA) % of hh in the community (EA) that received informal transfers Value of informal transfers (hh level)
	Disaster preparedness and mitigation	Additive Index (0-11)	Factor Analysis Index
		Village has improved infrastructure to mitigate shock impact (0-3) Disaster early warning system in place (0-3) People in village are aware of local coping mechanisms, disaster contingencies, and "risk resource maps" (0-3) The village VDC organizes DRR awareness events annually (0-2)	Presence of health center in the community Members in community migrate seasonally

Table D1: Original and own variable selection for estimation of TANGO method

		TANGO (Smith and Frankenberger, 2018)	Adaptation to LSMS-ISA data
	Cash Savings	Total value of current savings	N/A (information just available on third wave)
Adaptive Capacity	Bridging Social Capital	Factor Analysis Index	Factor Analysis Index
		Regular communication with people outside of one's village Engagement in economic activities with people from other villages Travel outside of one's village Number of VDCs of other villages that have shared learning	Presence of cooperatives in the community Share of males that participate in the coop. Share of females that participate in the coop. Value of informal transfers (hh level)
	Linking Social Capital	Factor Analysis Index	N/A
		Information on whether household members are friends or relatives of government officials	Dummy for whether a household member has works for government/political party
	Aspirations and confidence to adapt	Factor Analysis Index	N/A
		Absence of Fatalism Exposure to alternatives Women's freedom of movement Women's attitudes about family life	N/A (Information non available in Ethiopian LSMS-ISA data)
	Livelihood Diversity	Additive Index (0-11)	Additive Index (0-7)
		Total number of livelihoods the household engages in	Total number of activities that the household receives income from
	Human Capital	Factor Analysis Index	Factor Analysis Index
		Education of household members Literacy of household members Disability status of household members	Maximum level of education in the household Number of literate members Number of members in the household with impairing disability

Table D1: Original and own variable selection for estimation of TANGO method

		TANGO (Smith and Frankenberger, 2018)	Adaptation to LSMS-ISA data
Adaptive Capacity	Access to Information	Factor Analysis Index	Additive Index (0-4)
		Access to a cell phone Communication with people outside of one's village Number of program staff known Women's ability to travel to markets	Household ownership of: Fixed line telephone Mobile telephone Radio Television
	(*)		
Transformative Capacity	Access to markets	Walking distance to markets (village level)	Distance (Km) to the nearest large weekly market (village level)
	Access to services	N/A	Additive Index (0-7)
		% of households reporting access to each of 17 services (health care, family planning, primary school, preschool, gov council, dispute adjudication, and 11 others)	Presence in the community of: primary school, secondary school, health center, pharmacy, bus service, asphalted roads and agricultural extension agent
	Women's Empowerment	Factor Analysis Index	Factor Analysis Index
		Women's freedom of movement Degree to which women hold non-patriarchal values Women's decision making within their homes	At the village level: % of ag. hh where some/all land belongs to a woman Share of females that participate in coop. % of hh with a loan where the decision over it belongs to a woman % of the total hh with non-farm enterprises where business owned by a women % of hh that receive external income where its control belongs to a woman

Table D1: Original and own variable selection for estimation of TANGO method

		TANGO (Smith and Frankenberger, 2018)	Adaptation to LSMS-ISA data
		Factor Analysis Index	Factor Analysis Index
	Quality of governance	The capacities of VDCs (NGO-developed committees) Implementation of Community Action Plans Awareness regarding entitlements and responsiveness of government agencies	Needs for which community (EA) ask governmental institutions (out of 11) Needs for which governance consults community (EA) (out of 11) Mean level of needs addressed (EA level) (1=zero, 5=addressed)
	(**)		

NOTE: (*) Asset ownership also forms part of Adaptive Capacity; (**) Bridging and Linking Social Capital also contained in Transformative Capacity

Table D2: Original and own variable selection for estimation of RIMA

	RIMA-II (FAO, 2016)	**Own adaptation**
Access to Basic Services (ABS)		At the EA level:
	Inverse distance to non-agricultural market	Inverse distance to primary school
	Inverse distance to veterinary	Inverse distance to secondary school
	Inverse distance to primary school	Inverse distance to health center
	Inverse distance to health clinic	Inverse distance to bus service
	Inverse distance to input market	Inverse distance to paved road
	Infrastructure Index	Inverse distance to extension agent
		Inverse distance to pharmacy
		Inverse distance to periodic market
Assets (AST)	Per capita agricultural assets	Access to agricultural land
	Per capita wealth index	Durable asset index
	Per capita TLU	Livestock Ownership (TLU)
		Productive asset index
Social Safety Nets (SSN)	Transfers	Value of informal transfers (hh level)
	Other transfers	% of hh in the EA that received any type of assistance
	Scholarship (dummy)	Presence of cooperatives in the community (EA)
		% of hh in the EA that received informal transfers
Adaptive Capacity (AC)	hh average years of education	Max. level of education attainment (hh)
	Dependency ratio (inverse)	Num. of household members with disability
	Income generating activities	Number of distint income sources
	Inverse distance to input market	Access to information
	Inverse distance to non-agricultural market	
	Crop diversification index	

NOTE: the variables in each of the four pillars are combined trough factor analysis to generate the final pillar index

APPENDIX E

FACTOR LOADINGS OF RESILIENCE CAPACITY BUILDING VARIABLES

The generation of the resilience capacity variables proposed by TANGO entails the generation of multiple factor analysis indexes. Table E1 contains the factor loadings that yield those indexes. The rows in white correspond to the components of the factor analysis indexes of the rows in grey under which they are listed. The rows in grey are again combined through factor analysis to generate the final absorptive, adaptive and transformative capacity indexes. As we mentioned in previous sections of this document, for the sake of comparability, we include the resilience capacity building variables proposed by TANGO in the regression to generate the resilience score in the CB model. However, we estimate them through regular principal factor analysis, using all the three waves, not just the first wave to estimate the factor loadings, and without excluding variables due to ”wrong sign” or low KMO, as in the unconventional method used by TANGO. The ”Regular FA” in Table E1 therefore reports the factor loadings estimated using all the three waves and retaining the estimated factor loadings on all variables – letting the data speak for themselves – , while the ones under the ”TANGO FA” column are factor loadings estimated with just data from the first wave and with manual censoring of variables deemed to have factor loadings of the wrong sign or with a low KMO score. Consequently, the index values can differ between methods. Note that for the indexes that we will include in the CB, we do not group the resilience capacity building variables into absorptive, adaptive and transformative capacity indexes and therefore those factor loadings are not reported.

Table E1: Factor loadings, resilience capacity indexes (TANGO/CB)

	TANGO FA	**Regular FA**
Absorptive Capacity		
Bonding social capital	0.814	
Prob. of borrowing from close circle	0.205	0.301
% of hh receive informal transfers	0.338	0.405
Value informal transfers (birr)	0.258	0.252
Asset Index (0-100)	-	

	Factor Loadings	Factor Loadings
Durable Asset Index	-	-0.215
Productive Asset Index	0.558	0.512
Livestock (TLU) Asset Index	0.252	0.028
Access to ag. land	0.533	0.567
Access to informal SN	0.813	
% of hh that received assistance	0.317	0.352
Presence of cooperatives in the community	-	0.003
% of hh receive informal transfers	0.399	0.434
Value informal transfers (birr)	0.220	0.239
Disaster risk reduction	-	
Members in community migrate seasonally	-0.097	0.055
Presence of health center in the community	0.097	0.055
Adaptive Capacity		
Bridging social capital	0.264	
Value informal transfers (birr)	0.076	0.049
Presence of cooperatives in the community	0.727	0.693
Share of males participate coop.	0.742	0.734
Share of females participate coop.	0.696	0.734
Linking social capital	0.458	
Livelihood Diversity	0.108	
Human Capital	0.612	
Total years of education (hh)	0.691	0.691
Disability (hh)	-0.040	-0.032
Literacy	0.689	0.689
Access to Information	0.612	
Asset Index (0-100)	-	
Transformative Capacity		
Bridging social capital	0.491	
Linking social capital	0.160	

	Factor Loadings	Factor Loadings
Distance to main market	-0.127	
Access to services	0.485	
Women′s empowerment	0.329	
% of household enterprises owned by women	0.379	0.289
% of households where women decides over credit	0.314	0.346
% of households where women decides over income	0.406	0.299
% of households where women owns land	0.459	0.287
Share of females participate coop.	0.117	0.203
Quality of governance	0.226	
Needs comm. ask gov.	0.805	0.832
Needs gov. ask comm.	0.828	0.833
Mean level of needs addressed	0.357	0.072

NOTE:”TANGO” refers to the modified factor method and ”Regular FA” to standard principal factor.

Under TANGO, ”-” means that the variable was dropped due to ”wrong sign” or low KMO, (check Appendix D)

Table E2 shows the factor loadings for the generation of the four pillars in RIMA. The method used here is regular principal factor analysis. There is no manual variable selection according to sign or KMO; all the variables are included in the final index.

Table E2: Factor loadings to generate the four pillars (RIMA)

	RIMA
Access to Basic Services (ABS)	
Inverse distance to primary school	0.000
Inverse distance to secondary school	0.543
Inverse distance to health center	-0.027
Inverse distance to bus service	0.584
Inverse distance to paved road	0.339
Inverse distance to extension agent	-0.175
Inverse distance to pharmacy	0.389

	Factor Loadings
Inverse distance to periodic market	0.354
Assets (AST)	
Access to agricultural land	0.567
Durable Asset Index	-0.215
Livestock Ownership (TLU)	0.028
Productive Asset Index	0.512
Social Safety Nets (SSN)	
Value of informal transfers	0.239
% of hh that received assistance	0.352
Presence of coops. in community	0.004
% of hh that received informal transfers	0.434
Adaptive Capacity (AC)	
Max. level of education attainment (hh)	0.601
Num. of hh. members with disability	-0.082
Num. of distinct income sources	0.076
Access to information	0.613

APPENDIX F

CISSÉ AND BARRETT MAIN REGRESSIONS

Table F1: Results of regressions to generate mean of well-being p.d.f

	RAEC	**FCS**
RAEC, (t-1)	0.40***	
	(0.10)	
RAEC, (t-1), sq. (*10^4)	-0.05	
	(0.17)	
RAEC, (t-1), cub. (*10^8)	-0.07	
	(0.08)	
FCS, (t-1)		0.30*
		(0.18)
FCS, (t-1), sq. (*10^2)		-0.11
		(0.42)
FCS, (t-1), cub (*10^3)		0.01
		(0.03)
Drought	61.86	-0.25
	(68.08)	(1.19)
Bonding capital	-27.81	1.26
	(150.60)	(1.37)
Bridging capital	-50.84	-0.12
	(58.99)	(0.71)
Linking capital	465.6*	3.67*
	(276.50)	(2.16)
Asset Index	182.90	1.93**
	(115.50)	(0.95)
Livelihood Diversity	15.07	-1.859*
	(86.99)	(0.95)
Access to services	20.19	0.411
	(55.13)	(0.53)
Human capital	-40.22	0.682
	(63.38)	(0.806)
Access to information	362.5***	1.508***
	(54.00)	(0.49)
Women's empowerment	-141.40	0.05
	(120.60)	(1.34)
Quality of governance	-66.01	0.27
	(63.88)	(0.65)
Informal Safety Nets	168.70	2.84**
	(161.60)	(1.42)
Disaster risk reduction	-977.20	-27.56***
	(749.40)	(9.03)
Wealth Quintile	16.3***	1.16***
	(40.21)	(0.29)
Age of hh head	2.25	-0.0144
	(3.32)	(0.03)
Primary Education (head)	92.80	-0.049
	(70.69)	(0.82)
Secondary Education (head)	460.5***	1.078
	(114.0)	(1.23)

	RAEC	FCS
College Education (head)	1,223***	2.499
	(331.8)	(2.147)
Household Size	-161.0***	0.51**
	(21.25)	(0.21)
Percentage of males (0-4 y.o.)	-206.9	0.27
	(426.8)	(3.36)
Percentage of males (5-15 y.o.)	-520.7	-2.36
	(428.2)	(2.99)
Percentage of females (5-15 y.o.)	-132.7	-1.09
	(422.4)	(3.03)
Percentage of males (16-65 y.o.)	-161.0	-2.59
	(484.4)	(3.23)
Percentage of females (16-65 y.o.)	320.8	-3.64
	(421.5)	(3.05)
Percentage of males (>65 y.o.)	-165.1	-0.26
	(566.1)	(3.81)
Percentage of females (>65 y.o.)	-154.2	-1.90
	(462.5)	(3.47)
Constant	1,874***	23.93***
	(336.90)	(3.19)
Observations	6,137	6,460
R-squared	0.29	0.21

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table F2: Results of regressions to generate variance of well-being p.d.f

	res.(RAEC)	res.(FCS)
RAEC, (t-1) $(*10^3)$	0.15	
	(0.09)	
RAEC, (t-1), sq. $(*10^8)$	-0.06	
	(1.27)	
RAEC, (t-1), cub. $(**10^7)$	-0.01	
	(0.01)	
FCS, (t-1)		0.01
		(0.01)
FCS, (t-1), sq. $(*10^3)$		0.27
		(0.26)
FCS, (t-1), cub. $(*10^4)$		0.03
		(0.03)
Drought	-0.06	-0.04
	(0.08)	(0.05)
Bonding capital	-0.11	0.07
	(0.13)	(0.10)
Bridging capital	-0.03	0.05
	(0.04)	(0.04)
Linking capital	-0.34	0.05
	(0.33)	(0.09)
Asset index	-0.03	-0.15**
	(0.09)	(0.06)

Livelihood diversity	-0.06	-0.08
	(0.11)	(0.07)
Access to services	-0.05	-0.02
	(0.04)	(0.03)
Human capital	-0.12	-0.05
	(0.08)	(0.05)
Access to information	0.12**	0.06**
	(0.05)	(0.03)
Women's empowerment	-0.05	0.13**
	(0.08)	(0.06)
Quality of governance	-0.04	-0.02
	(0.05)	(0.04)
Informal safety nets	0.27*	-0.19**
	(0.14)	(0.10)
Disaster risk reduction	0.64	0.19
	(0.55)	(0.39)
Wealth Quintile	0.15***	-0.01
	(0.04)	(0.02)
Age of hh head	-0.01	-0.01
	(0.01)	(0.01)
Primary Education (head)	-0.11	0.03
	(0.08)	(0.06)
Secondary Education (head)	0.17	0.09
	(0.15)	(0.08)
College Education (head)	0.53	0.21
	(0.38)	(0.14)
Household Size	-0.07**	0.03**
	(0.03)	(0.01)
Percentage of males (0-4 y.o.)	0.28	0.16
	(0.45)	(0.22)
Percentage of males (5-15 y.o.)	0.04	-0.01
	(0.41)	(0.22)
Percentage of females (5-15 y.o.)	0.34	0.15
	(0.35)	(0.21)
Percentage of males (16-65 y.o.)	0.71*	0.27
	(0.39)	(0.21)
Percentage of females (16-65 y.o.)	0.61	0.04
	(0.48)	(0.22)
Percentage of males (>65 y.o.)	0.89	0.19
	(0.56)	(0.32)
Percentage of females (>65 y.o.)	0.43	0.12
	(0.49)	(0.25)
Constant	14.53***	5.63***
	(0.37)	(0.19)
Controls	Yes	Yes
Observations	6,137	6,460

Robust standard errors in parentheses. res.= residuals
*** p<0.01, ** p<0.05, * p<0.1

APPENDIX G

REGRESSION OF THE BUFFERING CAPACITY OF WELL-BEING

Table G1: Buffering capacity of the resilience capacity measures

	FCS	RAEC	FCS	RAEC	FCS	RAEC	FCS	RAEC
RCI [RIMA], (t-1)	0.14***	30.52***						
	(0.04)	(4.17)						
RCI [RIMA], (t-1)*Drought	0.09	1.39						
	(0.07)	(4.72)						
RCI [TANGO], (t-1)			0.12*	12.98				
			(0.08)	(8.67)				
RCI [TANGO], (t-1)*Drought			0.42***	2.86				
			(0.13)	(11.92)				
RS [CB, FCS], (t-1)					0.25***	22.11***		
					(0.04)	(4.64)		
RS [CB, FCS], (t-1)*Drought					-0.15***	1.30		
					(0.04)	(3.95)		
RS [CB, RAEC], (t-1)							0.14***	26.28***
							(0.04)	(3.28)
RS [CB, RAEC], (t-1)*Drought							0.07	1.98
							(0.07)	(4.14)
Drought	-12.17*	-747.10*	-21.23***	-813.40	9.97**	-617.00	-8.05	-632.10*
	(6.15)	(420.00)	(6.16)	(521.50)	(6.87)	(625.60)	(8.02)	(360.40)
Drought^2	-4.19*	-353.00**	-5.84***	-378.90**	0.51	-274.50	-2.42	–263.60*
	(2.28)	(175.40)	(2.03)	(187.70)	(2.46)	(199.40)	(2.62)	(135.40)
Wealth Quintile	1.51***	279.10***	1.74***	319.30***	0.37	221.50***	0.79	184.80***
	(0.36)	(43.41)	(0.34)	(46.33)	(0.44)	(50.93)	(0.50)	(52.38)
Age of hh head	-0.02	-3.12	-0.01	-0.65	0.01	2.06	0.01	2.25
	(0.03)	(3.77)	(0.03)	(3.75)	(0.04)	(4.55)	(0.04)	(4.94)
Primary Education (head)	-0.02	156.00*	0.67	272.50***	-0.69	224.00**	-0.76	115.9
	(1.02)	(85.62)	(1.06)	(91.88)	(1.08)	(99.10)	(1.18)	(95.96)
Secondary Education (head)	2.29	502.70***	3.88**	771.50***	0.39	698.80***	0.38	468.80**
	(1.53)	(146.10)	(1.67)	(143.40)	(1.69)	(189.30)	(1.76)	(184.40)
College Education (head)	3.98**	1,315.00***	6.89***	1,952.00***	3.59	2,159.00***	2.78	1,581.00***
	(1.95)	(307.00)	(2.27)	(317.90)	(2.39)	(493.40)	(2.81)	(506.30)
Household Size	0.93***	-204.10***	0.99***	-175.80***	0.33	-215.80***	1.40***	-92.40***
	(0.17)	(21.88)	(0.18)	(21.51)	(0.24)	(28.70)	(0.26)	(33.57)
Percentage of males (0-4 y.o.)	-2.27	-50.11	-2.06	47.03	4.16	152.80	2.86	-12.72
	(3.62)	(424.20)	(3.63)	(439.30)	(5.03)	(678.40)	(4.87)	(666.90)
Percentage of males (5-15 y.o.)	-4.91	-1,12***	-4.76	-992.70**	-2.86	-984.40*	-0.53	-491.20
	(3.32)	(392.00)	(3.41)	(408.00)	(4.42)	(569.30)	(4.53)	(591.80)
Percentage of females (5-15 y.o.)	-3.16	-476.80	-3.16	-398.30	1.74	-438.60	2.93	-161.50
	(3.50)	(411.60)	(3.49)	(421.80)	(4.22)	(645.80)	(4.52)	(640.50)
Percentage of males (16-65 y.o.)	-5.45	-314.30	-4.02	-30.41	-1.46	-494.40	-1.48	-318.10
	(3.84)	(468.30)	(3.94)	(478.70)	(4.89)	(610.00)	(4.84)	(617.20)
Percentage of females (16-65 y.o.)	-7.49**	-3.63	-6.78*	280.1	-0.65	568.90	-0.52	529.70
	(3.41)	(399.50)	(3.55)	(415.70)	(4.64)	(603.10)	(4.67)	(597.50)
Percentage of males (>65 y.o.)	1.48	323.20	1.96	468.60	-2.35	-404.70	2.15	-131.80
	(4.19)	(638.20)	(4.46)	(682.00)	(6.01)	(942.30)	(5.98)	(1,020.00)
Percentage of females (>65 y.o.)	-5.23	-363.10	-5.74	-275.30	0.28	18.10	-0.95	126.90
	(4.22)	(482.60)	(4.31)	(504.60)	(5.74)	(644.00)	(6.28)	(662.80)
Observations	7,274	7,046	6,768	6,556	3,187	3,082	3,074	2,996
R-squared	0.079	0.208	0.084	0.186	0.156	0.256	0.075	0.267

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

APPENDIX H

ROBUSTNESS CHECK: WINSORIZING THE WELL-BEING OUTCOMES

Since the distributions of our two outcome variables, the FCS and the RAEC are right-skewed, specially in the case of the RAEC, as a robustness check, I proceed to winsorize those variables symmetrically and re-estimate the resilience capacity measures and comparative tests. The main purpose is to determine if the distributions' outliers have a significant effect on the results. The winsorizing carried out is symmetrical, dropping the bottom and top 2% of the FCS and RAEC distributions. In this Appendix, I present the results of the re-estimation of the resilience capacity measures and the comparative results.

Figure H1 shows a comparison of the non-winsorized and winsorized FCS and RAEC distributions. The graphs on the left column are the FCS and RAEC distributions before winsorizing. The graphs on the right column show the FCS and RAEC distributions after eliminating the bottom and top 2% of those distributions.

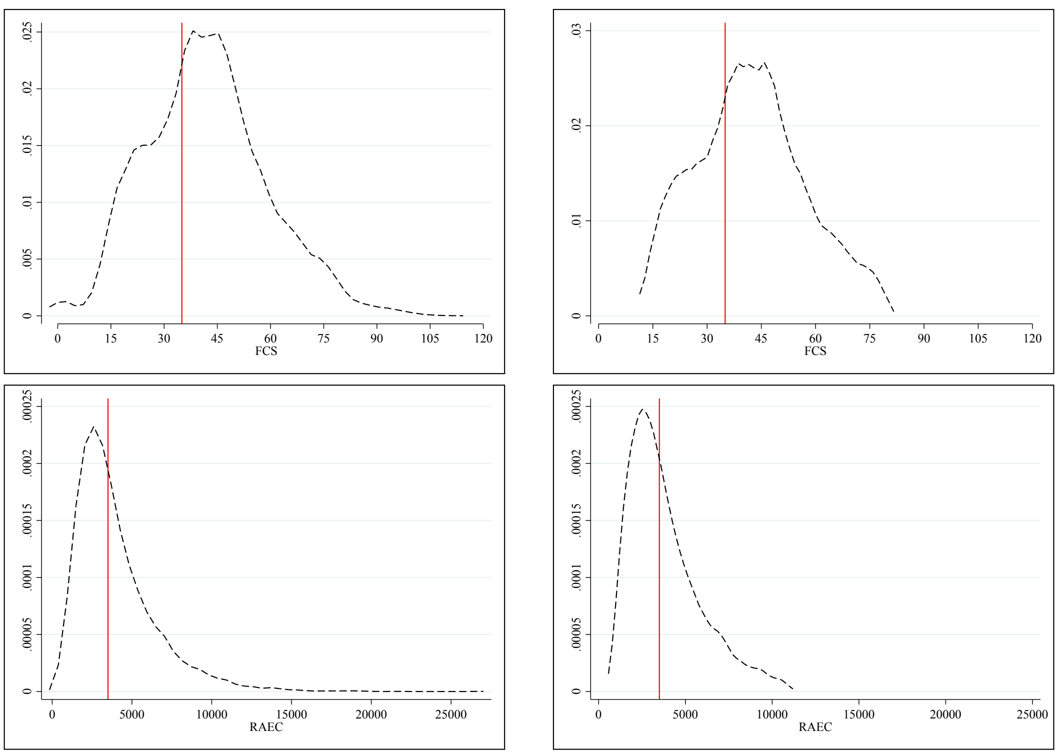

Figure H1: Original (left) and winsorized (right) FCS and RAEC distributions

Table H1 shows the summary statistics of the re-estimated resilience capacity measures. The initial findings remain practically unchanged. All the resilience capacity measures have been slightly scaled up, but the relationships between them remain the same. RIMA and TANGO RCIs are still very similar to one another, and CB's RS remain mostly unchanged.

Table H1: Summary statistics of the resilience measures after winsorizing

		Mean	SD	Min	Max
TANGO	Resilience Capacity Index (RCI)	22.10	10.65	0	100
RIMA	Resilience Capacity Index (RCI)	29.37	15.95	0	100
CB	Resilience Score [RAEC]	41.25	20.64	0.34	97.57
	Resilience Score [FCS]	68.03	15.51	21.62	99.29

Figure H2 shows the comparison of the non-parametric densities of the resilience capacity measures. The results remain very close to the original ones. RIMA and TANGO distributions appear again very similar. CB's resilience measures are now less skewed, their densities being more centered through all the range.

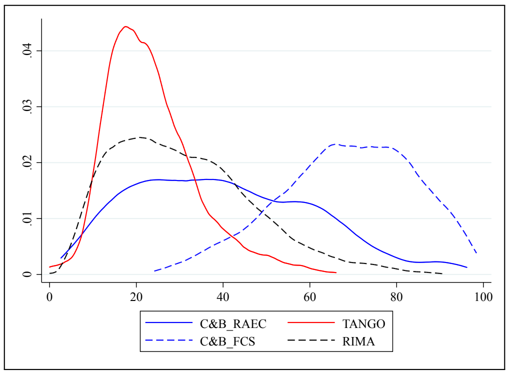

Figure H2: Comparison of non-parametric densities after winsorizing

Table H2 shows the results of the Spearman rank correlation coefficients between the resilience capacity measures and well-being outcomes. The results remain, again, practically unchanged. RIMA and TANGO's ranking by resilience capacity is highly correlated, and both are much less distant from CB's ranking. CB's measures are again the ones closest to standard well-being outcomes. Since the results of the ranking remain practically unchanged, we do not repeat here the comparison analysis of the bottom 20%.

Tables H3 and H4 show the results of the internal consistency checks. The resilient and non-resilient groups under each classification method were generated with the same procedure as in the main document. Under RIMA/TANGO, resilient households are the ones that do not suffer a loss in well-being between waves two and three. Under CB, resilient households are the ones that have an estimated conditional probability of being food secure higher than 65%,

Table H2: Spearman rank correlation coefficient matrix, after winsorizing

	RIMA	TANGO	CB[FCS]	CB[RAEC]	FCS
TANGO	0.683				
CB[FCS]	0.348	0.381			
CB[RAEC]	0.400	0.258	0.209		
FCS	0.178	0.193	0.361	0.076	
RAEC	0.215	0.141	0.165	0.493	0.286

NOTE: All the coefficients are statistically significant at the 0.01 level. Data from wave 3.

when the classification is done with the FCS; and those households that have an estimated conditional probability of non-being poor higher than 40%, when the classification is done with the consumption expenditures.

The results for the CB method remain unchanged. The resilience score and the binary classification of resilience are internally consistent, the group of households classified as resilient having a significantly higher average resilience score. In the case of the RIMA and TANGO methods, the results are not as robust. In both of them, the internal consistency between the RCI and the binary classification of resilience alternates depending on the well-being outcome with which the resilience classification has been done. RIMA appears to be consistent when the resilience classification is done with the consumption expenditures, and TANGO when the classification is done with the FCS. Nevertheless, even in those cases, the average values of resilient and non-resilient groups are very close to each other. To see how similar the resilient and non-resilient groups are, even when the t-test rejects the null of equality of means, I present in Figure H3 the comparison of the non-parametric densities of the RCI from the resilient and non-resilient groups that correspond to those t-tests with low p-values. The graph on the left shows the comparison of the non-parametric densities of TANGO's resilience capacity measure (RCI) for the resilient and non-resilient groups, classified by the RIMA/TANGO method with FCS as the well-being outcome of interest (data from wave 3). Those groups are the ones compared in the t-test that yields the p-value equal to 0.01 in table H4. The graph on the right shows the comparison of the non-parametric densities of RIMA's RCI for the resilient and non-resilient groups, classified by the RIMA/TANGO method with RAEC as the well-being outcome of interest (data from wave 3). Those groups are the ones compared in the t-test that yields the p-value equal to 0.06 in table H4. The graphs show how similar resilient and non-resilient groups are in terms of resilient capacity under the RIMA/TANGO binary classification method.

Regarding the differences in well-being outcomes between resilient and non-resilient groups, the results remain the same as in the original approach before winsorizing. Resilient and non-resilient households have significantly different values of FCS and RAEC under both CB's and RIMA/TANGO's binary resilience classification methods.

Table H3: t-tests (Resilience Score) between resilient/non-resilient groups, CB

	Resilience classification by: FCS			Resilience classification by: RAEC		
	Non-Res.	Res.	t-test (p-value)	Non-Res.	Res.	t-test (p-value)
CB[FCS]	52.68 (0.29)	78.71 (0.22)	0.00	66.08 (0.41)	71.20 (0.43)	0.00
CB[RAEC]	35.68 (0.57)	43.31 (0.54)	0.00	24.07 (0.26)	58.28 (0.37)	0.00
FCS	36.81 (0.41)	45.67 (0.36)	0.00	41.18 (0.39)	43.32 (0.41)	0.00
RAEC	2983.75 (50.55)	3400.61 (46.33)	0.00	2572.55 (35.28)	3972.14 (54.61)	0.00
Observations	1,027	1,596		1,377	1,246	

NOTE: In each row, mean values of RS and well-being measures for resilient/non-resilient groups. Standard errors in parenthesis.

Table H4: t-tests (RCIs) between resilient/non-resilient groups, TANGO/RIMA

	Resilience classification by: FCS			Resilience classification by: RAEC		
	Non-Res.	Res.	t-test (p-value)	Non-Res.	Res.	t-test (p-value)
RIMA[RCI]	30.98 (0.40)	30.97 (0.39)	0.97	30.52 (0.37)	31.56 (0.43)	0.06
TANGO[RCI]	22.91 (0.29)	23.97 (0.29)	0.01	23.43 (0.28)	23.52 (30)	0.84
FCS	36.37 (0.32)	47.61 (0.34)	0.00	40.83 (0.33)	44.61 (0.40)	0.00
RAEC	3106.86 (44.08)	3399.78 (44.19)	0.00	2634.02 (30.89)	4067.91 (52.51)	0.00
Observations	1,495	1,792		1,837	1,450	

NOTE: In each row, mean values of RCI and well-being measures for resilient/non-resilient groups. Standard errors in parenthesis.

Tables H5 and H6 show the results of the cross-sectional and panel well-being predictive accuracy of the resilience capacity measures after winsorizing the outcome variables. The original results are still maintained here. CB's resilience measures performance outstands the one of both RIMA and TANGO RCIs, both cross-sectionally and in panel. The results of the pair comparisons of the cross-sectional RMSEs (Table H7) are also maintained here. The distributions of cross-sectional RMSEs are different between measures, except in the case of the TANGO and RIMA well-being prediction . when FCS is the outcome of interest. Since the results of well-being prediction remain practically unchanged, and the well-being predictive test with the resilience capacity measures is the one with a higher empirical interest, we do not repeat here the additional approach of predicting well-being by each model's formal specification of the relationship between resilience capacity and well-being outcomes.

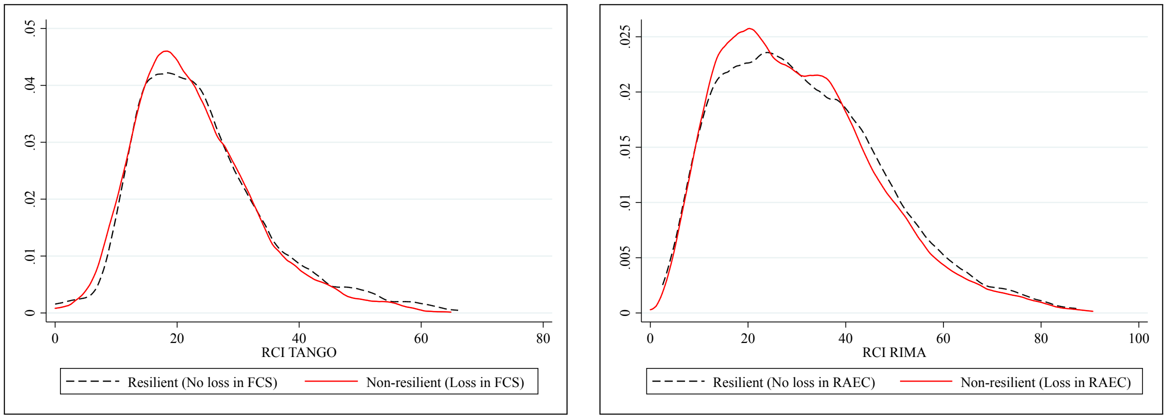

Figure H3: RCI's non-parametric densities of res./non-res. groups, after winsorizing

Table H5: RMSEs from cross-sectional well-being prediction, after winsorizing

	TANGO	**CB**	**RIMA**
FCS	14.28 (0.32)	13.59 (0.32)	14.28 (0.31)
RAEC	1714.60 (37.39)	1476.90 (34.32)	1667.33 (30.05)

NOTE: Standard errors in parenthesis. Estimation and prediction in wave 3

Table H6: RMSEs from panel well-being prediction, after winsorizing

	TANGO	**CB**	**RIMA**	**FCS_{t-1}**	**$RAEC_{t-1}$**
FCS	14.39	13.58	14.65	13.99	14.67
RAEC	1749.96	1554.35	1753.71	1766.16	1642.2

NOTE: Estimation in all-but-final survey waves, prediction in last wave

Table H7: p-values of pair comparison of cross-sectional RMSEs, after winsorizing

	TANGO-CB	**TANGO-RIMA**	**CB-RIMA**
FCS	0.00	0.95	0.00
RAEC	0.00	0.00	0.00

NOTE: Distribution of the RMSEs from 10 random draws of cross-sectional data.